DUPLICATE

Forecasting Research

Forecasting Research Division
Technical Report No. 186

GLOSS Parallel Trial Report

A Review of the GLOSS Parallel Run November 28 to December 18 1995

by

Brian Barwell and Stephen English

February 1996

Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

ORGS UKMO F
National Meteorological Library
FitzRoy Road, Exeter, Devon. EX1 3PB

Forecasting Research
Technical Report No. 186

GLOSS Parallel Trial Report

A Review of the GLOSS Parallel Run November 28 to December 18 1995

by

Brian Barwell and Stephen English

February 1996

Forecasting Research
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

© Crown Copyright 1996

This paper has not been peer-reviewed and should be regarded as an Internal Report from the Meteorological Office. Permission to quote from it should be obtained from the above Meteorological Office division.

GLOSS PARALLEL TRIAL REPORT

Abstract

Analysis of the GLOSS trial 28 November to 18 December has been completed. It was demonstrated that GLOSS data considerably reduces the r.m.s.e. of the forecast, particularly when verified against observations and most especially in the northern hemisphere and for the medium range. Key parameters for which a major improvement was seen are surface pressure, surface and jet level winds and surface temperature. This was achieved without significant degradation of other parameters. Whilst the benefit of GLOSS can be demonstrated statistically for most areas it can also be shown by study of specific meteorological features which have been more energetically forecast in the GLOSS trial run.

We recommend that GLOSS becomes part of the operational suite for the following reasons:

- GLOSS has an almost universally positive impact in the northern hemisphere at T+48 and beyond, and r.m.s. error is reduced by 10% (equivalent to 12 hours) at T+120.
- The impact of GLOSS in the tropics is positive for winds, pmsl, height and surface temperature.
- The Global Composite Index for NWP is estimated to increase.
- Synoptic studies suggest GLOSS impact is generally at least as good as the operational run but keeps details of smaller features better than the operational run.
- Experience at ECMWF and NMC using a similar retrieval procedure to GLOSS are very similar to the GLOSS results and these centres now use their GLOSS-equivalents operationally.
- GLOSS has a positive impact on the background wind fields, espeically at jet altitude.

Even where negative r.m.s.e. impacts are found, mostly in the southern hemisphere, experience at ECMWF and evidence from the GLOSS run suggests that the trial forecasts are more energetic and therefore capable of giving a better forecast than the operational run, even if this can't be demonstrated in r.m.s.e. There is also some evidence that GLOSS substantially improves the forecasting of tropical storm tracks.

Past documented problems with large unexplained bias over the sea ice have been overcome by a modification which has proved successful in the parallel trial, the re-run of the January 1994 period and monitoring of bias during the winter of 1995-96.

GLOSS PARALLEL TRIAL REPORT

A Review of the GLOSS Parallel Run November 28 to December 18 1995

by Brian Barwell and Stephen English

1. Introduction

Based on the results of the GLOSS parallel run Nov 1995 we recommend that GLOSS goes operational at the earliest possible opportunity. The GLOSS code for processing satellite data into a form which can be assimilated into NWP will replace the SAT120 and SAT500 data processed at NOAA-NESDIS. GLOSS has been tested in a parallel experiment running from 28 November to 18 December 1995 and also for a short period at the end of the 30 level model parallel run in January and February 1996. We have also carried out a re-run of the January 1994 case to check that the problems with GLOSS processing found for that period have been overcome. In this review we evaluate the impact in four ways: 1. Verification scores against observations and analysis in root mean square error (r.m.s.e.), bias and anomaly correlation; 2. Impact on background field (for windspeed); 3. Examples of meteorological synoptic features better defined in the trial run; 4. The NWP Composite Index (CI). All verification against analysis is for the operational analysis.

The verification scores are based only on December data as the new verification software could not cope with a change of month. In this trial the first four days of the run would be treated as a warm up period anyway, so we were not adversely affected, but for future trials verification software should be able to cope with data including a change of month.

Results from additional studies carried out by Roger Grant, Richard Graham and Julian Heming are included as annex 1, annex 2 and annex 3. The major impacts of GLOSS are summarised in the conclusions. A complete list of results is accessible from the Satellite Sounding Group homepage.

In this report we will demonstrate that GLOSS data considerably reduce the r.m.s.e. of the forecast, particularly when verified against observations and most especially in the northern hemisphere and for the medium range. Key parameters for which a major improvement was seen are surface pressure, surface and jet level winds and surface temperature. This was achieved without significant degradation of other parameters. Whilst the benefit of GLOSS can be demonstrated statistically for most areas it can also be shown by study of specific meteorological features which have been better forecast in the GLOSS trial run. Examples are given.

As well as providing a measurable statistical improvement in performance it has also been demonstrated that the trial forecasts were more energetic i.e. they captured small synoptic features more convincingly than the operational run. This is consistent with the experience of other centres who have demonstrated positive impacts and are already assimilating satellite data operationally using methods very similar to GLOSS. In particular we shall compare with the experience at ECMWF where relevant in this review.

2. Verification scores: Analysis by region

2.1 Northern Hemisphere from 30N to 90N

The GLOSS run showed a growing positive impact in Northern Hemisphere (NH) mean sea level pressure (pmsl) after T+24, and a neutral impact up to T+24 (Fig 1). In all figures in this report the continuous line is the operational (control) run and the dashed line is the trial (GLOSS) run. By T+72 the improvement corresponded to 6 hours in the forecast period, or about a 7% reduction in r.m.s.e. In Fig. 1 the error in the GLOSS run is observed to grow less quickly than in the operational run such that the biggest positive impact was at T+120. This result is consistent with the results from the November 1993 parallel run (Gadd *et al.* 1995), despite a much lower r.m.s.e. error in the 1995 operational run. Comparison of the verification of pmsl against observations and analysis showed no significant differences: in both, the GLOSS run gave a broad positive impact. The trend of the error to grow less quickly in the GLOSS run was observed in both verifications. There was a corresponding positive impact in surface winds, with the errors correlating very strongly with those from pmsl. Here too the impact was neutral at first against observations or slightly negative against analysis to T+24 but thereafter a growing positive impact was observed. The positive impact for medium range products was very considerable rising to over 5% in r.m.s.e. by T+72, as can be seen from Fig. 1 for NH pmsl.

The height fields also showed better verification against observations. The NH height fields showed a neutral impact to T+48 (slightly negative against analysis) but thereafter a growing positive impact, reaching over 5% by T+96. This is a marked improvement on previous GLOSS trials. The verification at T+96 is given as an example in Fig. 2. The negative impacts are very small. The wind profile showed neutral impact until T+96 with a small positive impact at T+120, although the verification against radiosonde winds (as opposed to aireps) was markedly better. Generally positive impacts were observed at jet altitude.

Richard Graham repeated the study used in Gadd *et al.* (1995) showing the vector differences from the background field for the operational run and the GLOSS trial for different areas and windspeed categories (at all levels). In the NH, defined here as 45N to 90N, positive impact was seen for windspeed greater than about 20m/s, and the positive impact was large for windspeed greater than 45m/s (jet level winds). The results from Graham's study are included in Annex 1. There is little doubt therefore that a major likely benefit of GLOSS data will be an improved windfield, particularly at low altitude and jet altitude. As the satellite does not directly sense windspeed this considerable benefit is most likely a result of an overall improvement in the consistency of other parameter fields.

Verification against humidity showed only a marginal positive impact, and this was late in the forecast run. However the bias, which is not much discussed for most parameters in this review, was noticeably reduced and no serious negative impact was found. The reduction of humidity bias is a very good result.

Annex 4?

No obvious error was found in the NH correlating with sea ice. The change following the last GLOSS trial was therefore successful (as errors over sea ice caused errors to grow in past trials). Nonetheless a re-run of the January 1994 case was performed in order to check the widespread bias problem had been overcome. The results were much improved in the re-run, although the

run can not be performed to the same standard as the current GLOSS as some key parameters (e.g. satellite scan position) are not available in the archived dataset. Overall the re-run results were very encouraging, as they were consistently improved and were strongly positive in the tropics and for temperature and wind fields, and to a lesser extent the humidity field.

Another way to look at forecast scores other than bias or r.m.s.e. is the anomaly correlation (AC). That is how well correlated are anomalies in the field of a given parameter by reference to weighted mean monthly climatology. Perfect correlation is given by an AC=1.0 and no correlation is indicated by AC=0.0. A good forecast will generally have a high AC.

The AC results were most encouraging. Fig. 3 shows that for mean sea level pressure the AC improvement is consistent with the improvement in r.m.s.e., with the increase in correlation of the trial run over the operational run growing from T+48 onwards to the equivalent of about 12 hours forecast time by T+120. In Fig. 4 the NH AC is given for height at 850hPa. The AC at 250hPa and 500hPa (not shown) are almost identical to that at 850hPa. Comparison shows that the AC improves at all heights, but that the improvement is seen earlier in the forecast period at 850hPa. In none of the NH height ACs is a negative impact seen at any time.

For temperature fields the AC did show a small negative impact at T+24 (except at 250hPa) but this turned into a positive impact at 500hPa and 850hPa. At 250hPa the AC impact was neutral throughout the forecast period.

2.2 Verification scores for the LAM domain (Area 2)

The NH results showed general positive impact. A study was also made of the impact on a more regional scale, although using UGM forecasts as a LAM trial did not take place. Firstly for an area corresponding roughly to the LAM domain and secondly just for the UK and its coastal waters. It should be noted that a two week trial will become progressively less representative for a smaller region, owing to persistence of weather types.

The verification scores in the LAM area were even better than those in the NH as a whole. A larger positive impact was seen in pmsl from T+48 to T+96 (over 10% at T+72). As with the whole NH the error grew less quickly in the GLOSS run such that the gap between GLOSS and operational became wider during the forecast period.

For all parameters the positive impact was larger and seen earlier in the forecast for the LAM area than the whole of the NH. In Fig. 5 the pmsl r.m.s.e. is plotted for the verification against observations. Comparison with Fig. 1 shows that the reduction in r.m.s.e. is both earlier and larger for this local area, being about 10% from T+48 onwards. On this occasion the verification against the operational analysis update was as good as that against the observations. A similar positive improvement was seen for all other parameters of interest.

The AC for height, temperature and pmsl are also available for area 2. These are not shown but the results were consistent with the verification using r.m.s.e.: a very large positive impact developed in pmsl from T+48 onwards, with a neutral or smaller positive impact at T+24 for temperature fields..

The r.m.s.e. was also noted to be substantially reduced and AC improved in the area immediately around the UK. Although encouraging, the representivity considerations mentioned above limit the weight which can be placed on this result.

2.3 The tropics (30S to 30N)

In the tropics the most important parameters are windspeed and temperature. Past GLOSS trials have shown a significant positive impact on the tropical winds. The verification results confirmed that GLOSS is strongly beneficial in the tropics, although the positive impact was slightly less marked than previously. There were still major benefits in the tropical windfield, but the temperature results were slightly less convincing. Richard Graham's analysis (see Annex 1 for full results) showed that the GLOSS data made a significant beneficial effect on the background windspeed compared to radiosondes. This was especially so at high windspeed, and at jet altitude, albeit based on an unsatisfactory number of data points. But at 20-30m/s there were sufficient verifications to have reasonable confidence in the statistics and here too a large positive impact was observed.

In the forecast verifications a more or less neutral wind impact was seen at high levels at T+24 and T+48, becoming positive by T+72. The verification against radiosonde winds at T+72 is given in Fig. 6, where it can be observed the bias is also reduced. At T+96 and T+120 a positive impact was seen. At low altitude small negative impacts against analysis were found out to T+48, thereafter the impact was mostly neutral. Fig. 6 shows that whilst the GLOSS data do not provide additional wind field information at lower altitude in a 1DVAR type retrieval, there is no detrimental impact on the lower altitude windfield, and the benefits at jet altitude are considerable.

The forecast verification against radiosondes showed a positive impact at high altitude from T+24 which became larger and extended to lower altitude through the forecast period. Aireps gave a neutral impact comparison for short range forecast times, although here too the impact was clearly positive by T+72 as is seen in Fig. 7. The bias is also reduced and there is encouraging consistency in these results from two independent verification data sources (comparing Fig. 7 with Fig.6 at 250hPa). The AC in the tropics results in the tropics varied widely, but AC has little meaning in this area.

Julian Heming studied the impact on the forecasting of tropical storms. His summary and examples are given in annex 3. Although only two tropical storms were reported during the trial period his analysis showed that in both cases the storm tracks were far better forecast in the GLOSS run. In fact the improvement was as good as the best cases following recent changes specifically targeting the forecasting of storm tracks, and far in excess of the improvement seen from any other parallel run in recent years. Therefore, although two storms are not enough to make statistically sound conclusions, it appears highly likely that another benefit of GLOSS would be substantially improved forecasts of storm tracks in the tropics.

2.5 Southern Hemisphere 30S to 90S.

The results already described for the NH, including details for the UK region, and the tropics showed very encouraging verification except for the tropical temperature field. By contrast the results from the Southern Hemisphere (SH) were less encouraging. They were however

consistent with results from the ECMWF, which nonetheless adopted operational assimilation of their equivalent of GLOSS. They argued there are reasons to expect a negative impact in the SH. This is discussed below. Negative r.m.s.e. impacts were seen both for height and pmsl. The bias was markedly reduced, as can be observed in both Fig. 8 which plots verification against observations for pmsl and Fig. 9 which shows the verification for surface windspeed. The r.m.s.e. did not grow very quickly with time, and for some parameters r.m.s.e. grew less quickly in the trial run than the operational run after T+48: e.g. for windspeed (see Fig. 8) where the difference in r.m.s.e. was smaller at T+120 than at T+72. This was however less frequently observed in the SH than the NH. Fig. 8 can be compared to Figs. 1 and 5 which show the same verification in the NH.

The AC in the SH showed results that were more encouraging than the r.m.s.e. verification. Although the impact at T+24 was again negative at all heights, the impact quite rapidly moved towards neutral during the forecast period, as Fig. 10 illustrates for temperature at 850 hPa. The negative impact at T+24 decreases with time becoming neutral by T+72 and remaining neutral thereafter. At 250 hPa the AC reached neutrality later, not fully eliminating the negative impact until T+120, although it was very slight after T+72.

The SH AC gave an indication of a mis-match of the analysis used in the verification by extrapolating the AC to time zero. This was done for temperature 850 hPa AC for the SH (the data shown in Fig. 10). The operational line extrapolates to unity at time zero, whereas the trial line extrapolates to less than 0.9. This is evidence that there is indeed a mismatching of the analysis by use of the operational analysis update for verification. The operational analysis is tuned to SAT120 biases. At the ECMWF, where the verification was against the trial's own analysis, a positive impact was seen in AC, despite a negative impact in r.m.s.e. very similar to that from GLOSS. At ECMWF it was found that the analysis has higher energy, and this resulted in better scores in AC than in r.m.s.e. The SH r.m.s.e. and AC are therefore misleading in this study. A general conclusion of this trial is that GLOSS should be validated against the trial analysis, although in this case it is not too serious a problem as the major benefits of the GLOSS data are still strong enough to be apparent in any verification except in regions of very strong flow in the SH.

The change in impact with time for AC for 250/500/850hPa height was also more encouraging in the SH than the r.m.s.e. verification. Overall the height AC for the SH, although mostly negative gave a better result than the r.m.s.e., and certainly the negative impacts here were much smaller than the NH positive impacts, which is consistent with results from ECMWF.

However statistical tests are only one way of testing the impact of GLOSS data on NWP and the value of GLOSS in the SH is discussed again in section 4 and in Annex 2 and Annex 3.

3. Global verification scores

3.1 Number of positive r.m.s.e. impacts for GLOSS trial

The verification against observations and analysis has been analysed to find the total number of clearly positive impacts, clearly negative impacts and broadly neutral impacts. To do this the following parameters were considered at T+24, T+48, T+72, T+96 and T+120: mean sea level pressure, surface pressure, surface temperature, surface windspeed, 500hPa height, 500hPa temperature, 850hPa humidity, windspeed at 250hPa and 850hPa. Because r.m.s.e. temperature errors at different heights were generally well correlated, the 500hPa temperature was taken to be widely representative of the whole temperature profile. Similarly the 850hPa level was taken as representative for humidity. It was felt that a windspeed was required at 250hPa and 850hPa as both these heights are important and the errors were not always well correlated. As was discussed in section 2, the short term comparison will be negative against the operational analysis so for this comparison against analysis the T+24 comparison is neglected. Broadly neutral is defined as a difference in forecast skill equivalent to less than 2 hours in the forecast period (usually about 1% in r.m.s.e). This analysis does not do justice to how large some of the positive impacts were in the NH, but is intended to give a quick guide to where the principal impact lies. Results are shown in Table 1.

Note that as positive impacts are difficult to achieve against the operational analysis update, the verification against observations is more meaningful, and this report has concentrated on this. In the past verification for GLOSS trial runs against the trial run,s own analysis has tended to be better than verification against observations. In general, impacts which are seen as neutral in the verification against observations are seen as marginally negative in the verification against the operational analysis, but the negative impact usually reduces with forecast range and often turns positive by T+72 in these cases. Table 1 clearly shows that, globally and in verification against observations, neutral or positive impacts (97) are more likely than negative impacts (23). In the NH the verification is very good. Negative impacts are rare (only 2 against observations and 5 against analysis). In the tropics the observations give good verification. In the SH the verification is generally disappointing. Reasons for this low level of impact in the SH were discussed in section 2.

Table 1: Number of r.m.s.e. positive/neutral/negative impacts by region and type

Region or data type	Verification against observations			Verification against analysis		
	positive	neutral	negative	positive	neutral	negative
NH	17	21	2	23	4	5
Tropics	18	21	1	10	7	15
SH	2	18	20	0	0	32
Profile	17	44	14	19	6	35
Surface	20	16	9	14	5	17
Global	37	60	23	33	11	52

The impact was different for surface parameters (pressure, windspeed, temperature) compared to profile variables (height, temperature, windspeed and humidity). The major positive impact of the GLOSS data is for surface parameters and jet level winds, though a broad spread of smaller positive impacts was seen for other parameters. Surface pressure is particularly improved in the trial run.

3.2 The NWP composite index

The NWP composite index (CI) is a normalised weighted mean of r.m.s.e. for 22 selected parameters and is given by:

$$CI = 100 \frac{\left[\sum_i W_i \frac{E_i}{\varepsilon_i}\right]}{\sum_i W_i}$$

where i denotes the ith parameter, W_i its weight, E_i is its r.m.s.e. value for 1991 and ε_i is the current r.m.s.e. (In practice, a 12-month running mean is taken with different E values for each month.) Using this formula with j as the period of the parallel trial and denoting r.m.s.e. values for the operational and parallel runs by ε_{0i} and ε_i, we can define the composite index for operational and trial run as CI_0 and CI respectively. To eliminate E_i we assume the ratio of operational r.m.s.e. is independent of this parameter which leads to:

$$\frac{CI - CI_0}{CI_0} = -\frac{\left[\sum_i \frac{W_i(\varepsilon_i - \varepsilon_{0i})}{\varepsilon_i}\right]}{\sum_i W_i}$$

In effect, this defines a relative change in composite index for the trial assuming that the improvements in trial and operational r.m.s.e. values over 1991 values for the trial period are representative of annual mean values.

This relative composite index for NWP was calculated for the GLOSS run with respect to the control run, taken from verification against surface and radiosonde observations. A description of CI weights can be found on the FS4 internal web homepage.

The global relative score was 0.0030, i.e. based on this trial we estimate a very marginal increase in CI of 0.3%. The relative CI (RCI) does not need to be calculated globally. It could equally well be applied by regions. This reveals that the RCI for the northern hemisphere is 0.0154, for the tropics 0.0166, but for the southern hemisphere -0.0296. It is apparent that from a purely statistical sense the GLOSS impact is very significant in the NH in terms of parameters of direct importance to our customers. In the tropics it is also positive, and the overall global impact is positive.

To estimate the relative composite index using verification against analyses, the relevant fields were extracted from archived fieldsfiles so that r.m.s.e. values for the trial could be computed using trial analyses for verification. Very similar results were found, with positive contributions

from the northern hemisphere and tropics offset by negative ones from the southern hemisphere. The global score was effectively neutral (-0.0003) and the regional values were: NH 0.0160, tropics 0.0042, SH -0.0352.

3.6 Summary

For the NH GLOSS gave an improvement in the forecast accuracy equivalent to 6 to 8 hours in the forecast period after T+72. for pmsl. The improvement was over 8 hours for the European area (Area 2) and over 10 hours for pmsl in the UK area. The r.m.s.e. results were backed up by substantial improvements in the AC. The results from the tropics were mixed, with positive impact on height and windfields but a negative impact on temperature field. In the SH the impact was undoubtedly negative in a statistical sense, but this was smaller and less significant than the positive impact in the NH. The bias has not been discussed much but generally it was reduced for most parameters most of the time. It was found that, for those areas which showed a negative impact with respect to the analysis, the AC suggested a mis-matching between the operational update analysis used in all verifications and the trial analysis, as the AC would not extrapolate to unity at time zero. Therefore there is evidence, but not proof, that these negative impacts were caused by a choice of analysis for verification not appropriate for GLOSS. The next section, for which we are grateful to Roger Grant for doing the analysis, takes a synoptic rather than statistical view of the impact.

4. Verification by synoptic features

Statistics can be unduly influenced by large differences between bad and very bad forecasts or large differences which have little effect on the weather (e.g. central pressure of the Siberian high). Any statistical analysis should be backed by a study of the capturing of synoptic scale features in the analysis. Roger Grant carried out a detailed study of features in the GLOSS and operational fields. His summary is included as Annex 2. His main conclusions were:

1. On most occasions GLOSS seemed to be able to keep details of the smaller scale better than the operational model, for example upper troughs (note: this can actually increase the r.m.s.e. if the feature is slightly misplaced, especially in strong flow.). We do not know if the trial was better as no validation was done, but trial forecasts were certainly more energetic which is in itself a good and positive result.

2. GLOSS forecasts were as good as the operational forecasts.

3. Subjective judgement showed no significant improvement of GLOSS over the operational forecasts.

4. GLOSS made surprisingly little difference to the SH analysis.

These conclusions whilst mixed, give some indication as to why the impact of GLOSS in the SH was poorer than in the NH. Firstly the background covariance matrix used isthe same as in the NH, so we would not expect the background field to be changed any more than in the NH. The ECMWF increase the background error standard deviation ssouth of 60S in a linear way; a feature which is may be introduced to GLOSS at a later date.

This would increase the weight given to the satellite data and of course increase the impact of GLOSS on the SH analysis. Secondly in the fast moving, land free, circulation around Antarctica it is easy to misplace a feature in space by significant distances whilst still producing what is essentially a good forecast (i.e. the timing of the movements may still be good). If the GLOSS data make the fields more energetic and sharpen the contrast of a feature from the background, as we know it does at ECMWF and is implied in Annex 2, then small timing errors will give a greater r.m.s.e. error. The higher r.m.s.e. in the SH does not necessarily imply an inferior forecast. Further information is needed and this can best be supplied by detailed study of a number of case studies, based perhaps on the examples given in Roger Grant's study. However a negative impact on r.m.s.e.,whilst not implying an inferior forecast, does lower the CI scoreand the overall effect of GLOSS in the SH is to almost eliminate the good scores achieved in the NH.

5. Conclusions

The parallel trial of GLOSS in November and December was successful. Large positive impacts were seen in many areas, mostnotably mean sea level pressure, surface and jet level wind fields, and most parameters at medium forecast range (at and beyond T+72). The Anomaly Correlation was markedly closer to unity for height and temperature fields in the Northern Hemisphere in the trial run. In the tropics the results were mostly good, but a negative impact was seen for temperature both in r.m.s.e. and AC. However the tropical surface temperature forecast was improved. It was difficult to demonstrate a positive impact in the SH. The AC suggested the verification against analysis was misleading and verification against observations is sparse (an order of magnitude less than in the NH). Synoptic studies suggested that the GLOSS data were producing forecasts at least as good as and more energetic than the operational run, but this can on occasion penalise the r.m.s.e. error in the forecast. A particularly large positive impact was seen for the North Atlantic and Western Europe, most especially around the British Isles. No evidence of the problemswhich caused the withdrawal of GLOSS during the previous attempted operational implementation were observed and we propose the GLOSS data to WGOS as a positive contribution to the NWP effort in the Office.

To summarise, the key conclusions are:

1. GLOSS has an almost universally positive impact in the NH at T+48 and beyond
2. GLOSS impact in the tropics is positive for winds, pmsl and height and surface temperature
3. GLOSS impactin the tropics is negative for the temperature profile
4. GLOSS impact on r.m.s.e. scores in the SH is negative although this is often reduced at longest forecast range
5. GLOSS impact is particularly beneficial in the region of the UK
6. GLOSS has an almost universally neutral impact according to these statistics at T+24
7. The Global Composite Index for NWP is estimated to increase by 0.3%
8. Synoptic studies suggest GLOSS impact in the SH may be more beneficial than statistics suggest

9. There is much higher confidence in the positive impacts in the NH and the tropics, which are supported by the AC and/or large numbers of observations than the negative impact in the SH which produce a more energetic flow but are supported by very few observations.

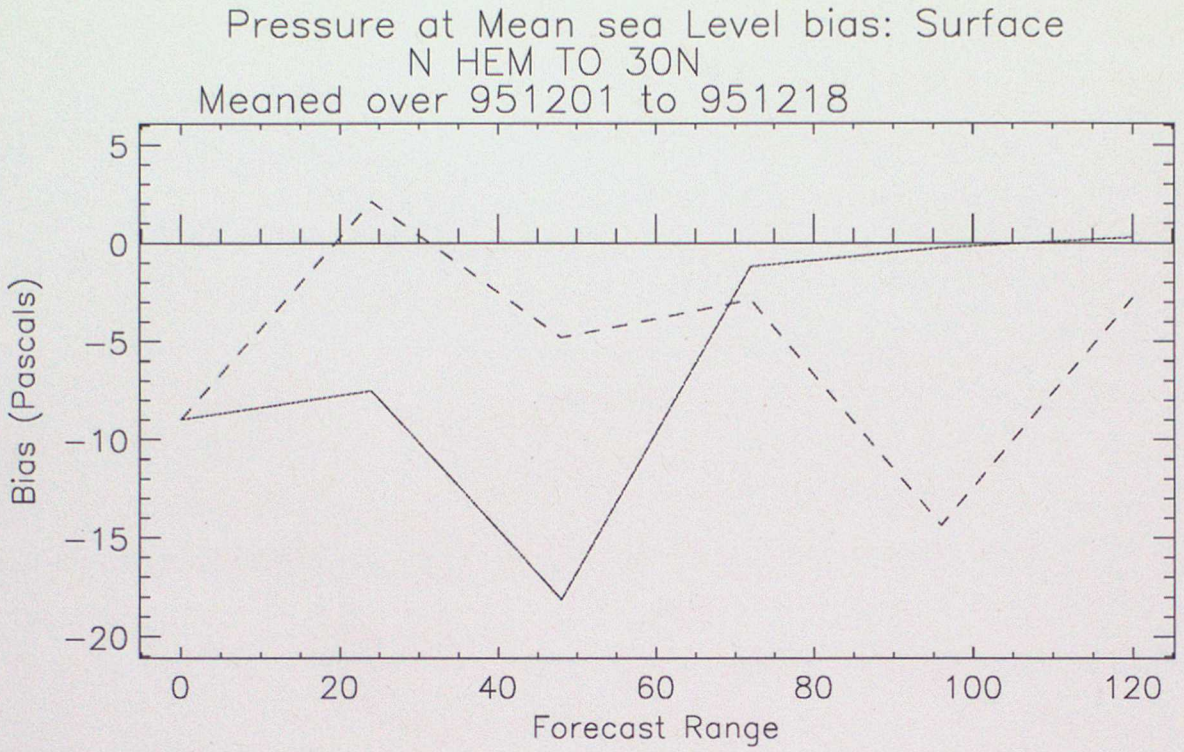
Annex 1: Richard Graham's study of impact on wind fields.

Annex 2: Roger Grant's synoptic study of GLOSS impact.

Annex 3: Julian Heming's analysis of tropical storm tracks.

6. Reference

Gadd A.J., B.R. Barwell, S.J. Cox and R.J. Renshaw 1995: Global processing of satellite sounding radiances in a numerical weather prediction system., Q. J. Royal Meteorol. Soc., **121**, 615-630.



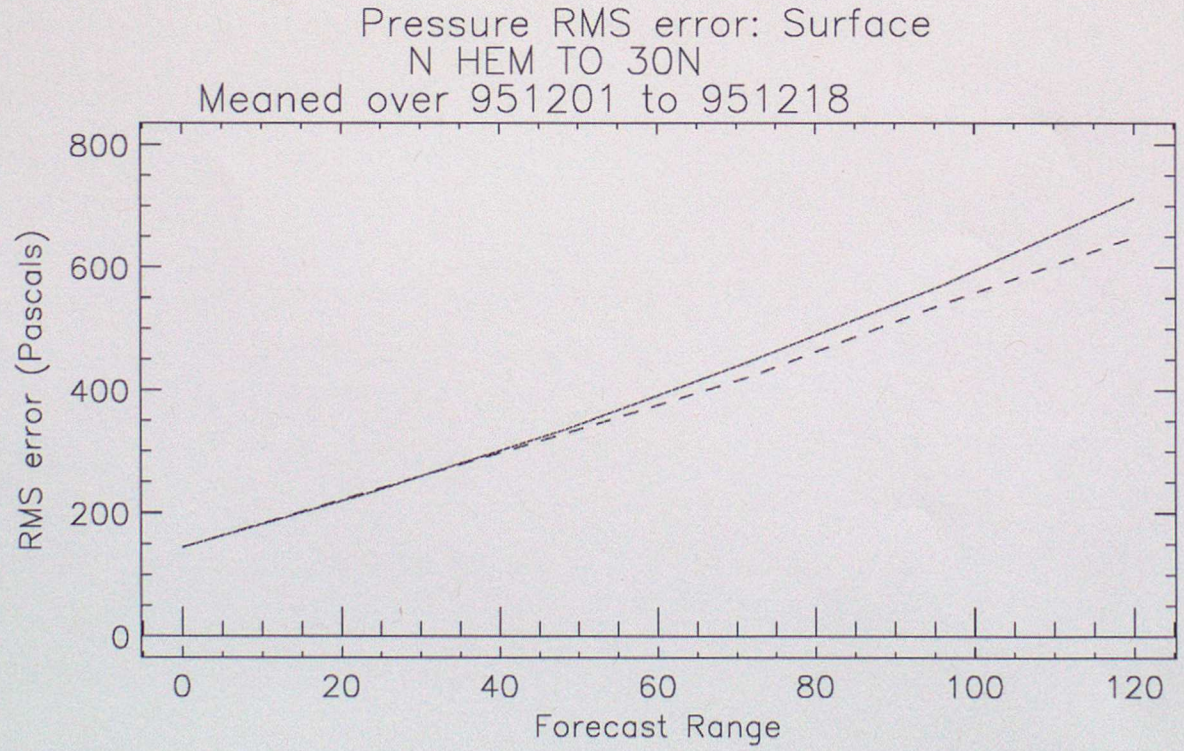
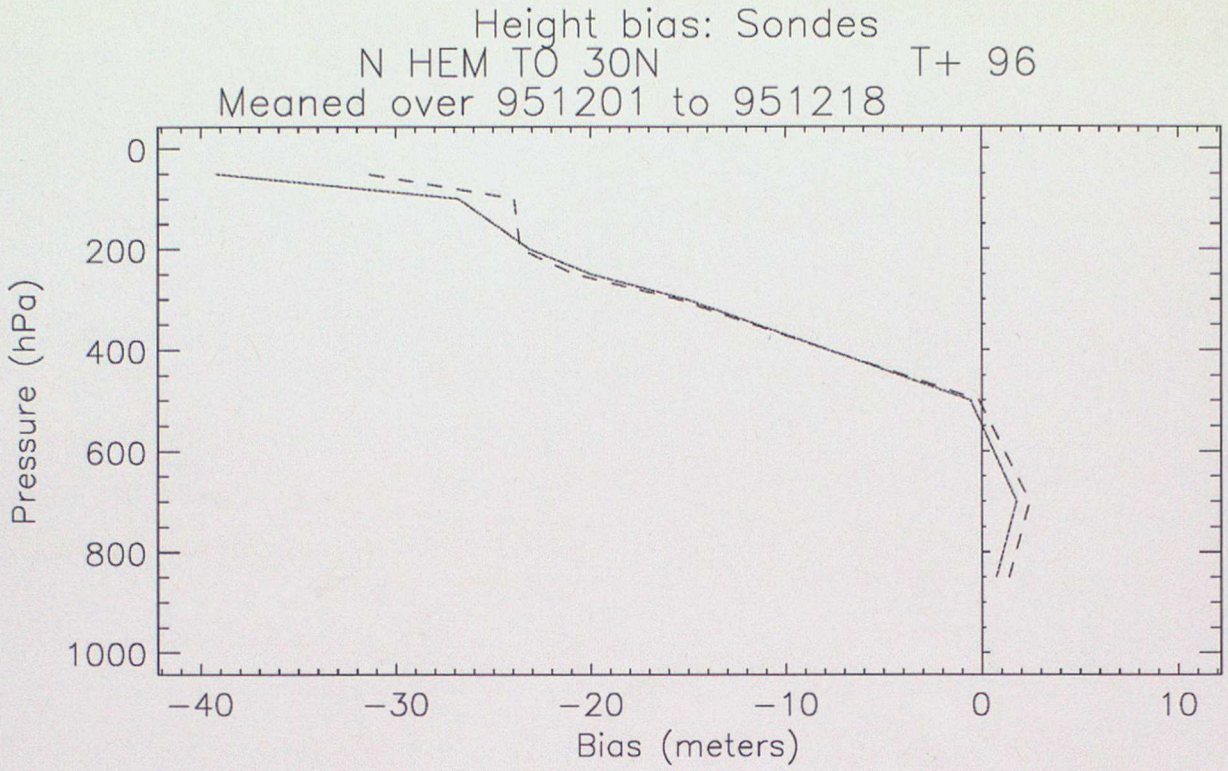


Fig. 1 NH mslp verification against observations. Continuous line denotes operational run and dashed line GLOSS run (this is retained for all Figs).



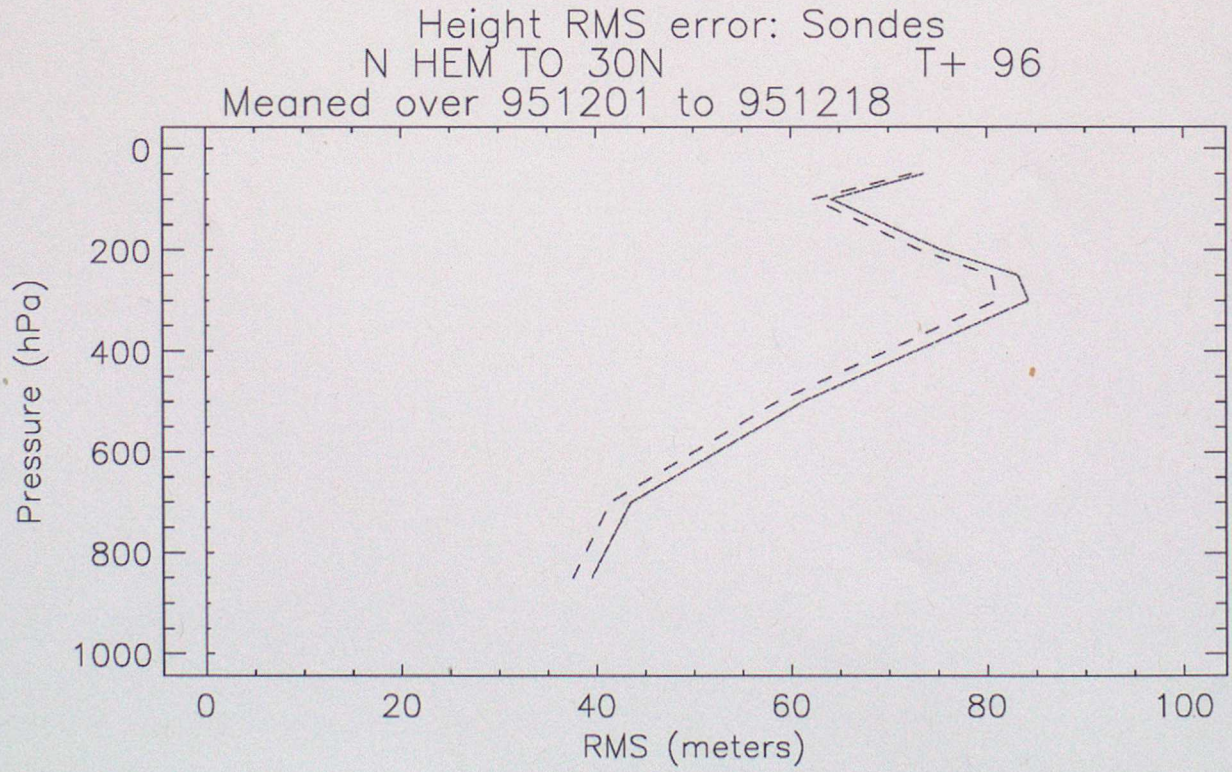


Fig. 2 NH height at T+96 verification against observations (lines as Fig. 1)

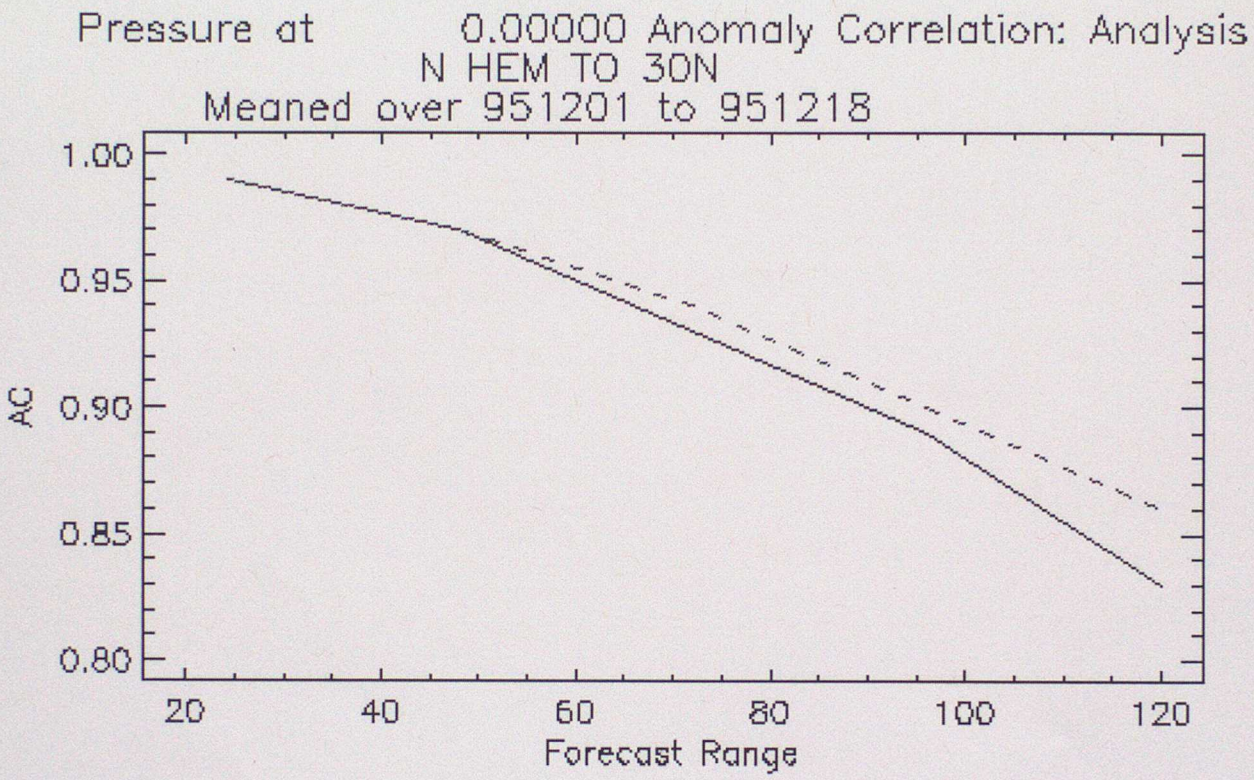


Fig. 3 NH Anomaly correlation (AC) for pmsl (lines as Fig. 1)

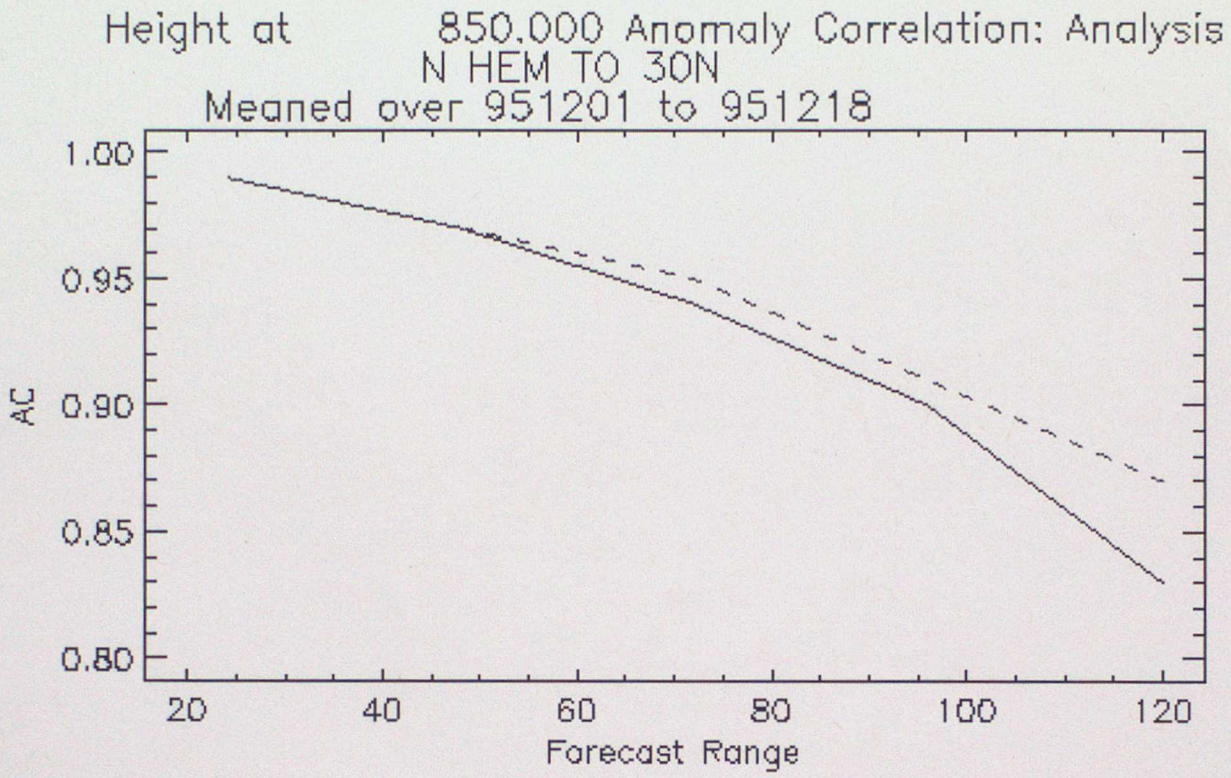
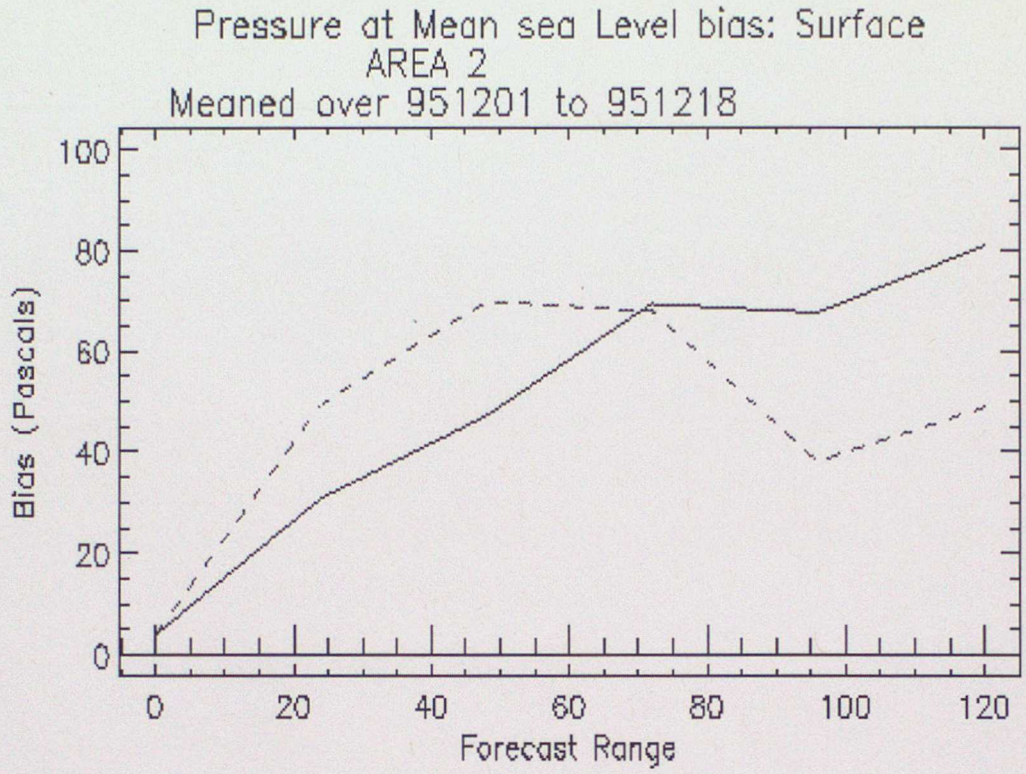


Fig. 4 NH Anomaly correlation (AC) for 850hPa height (lines as Fig. 1)



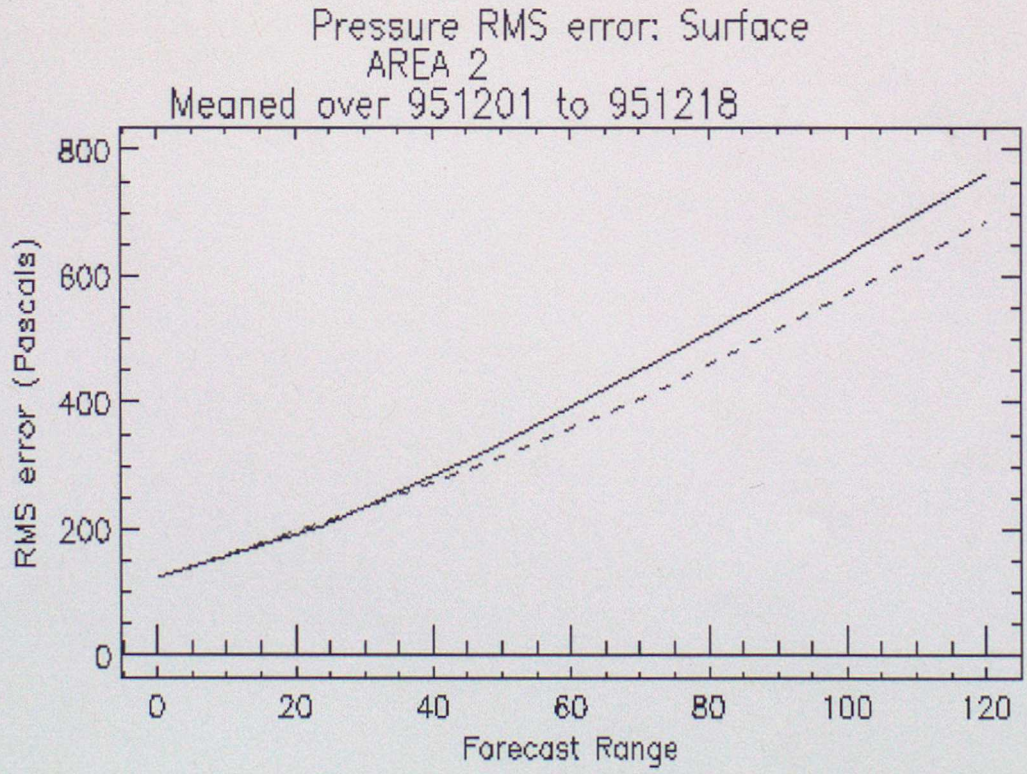
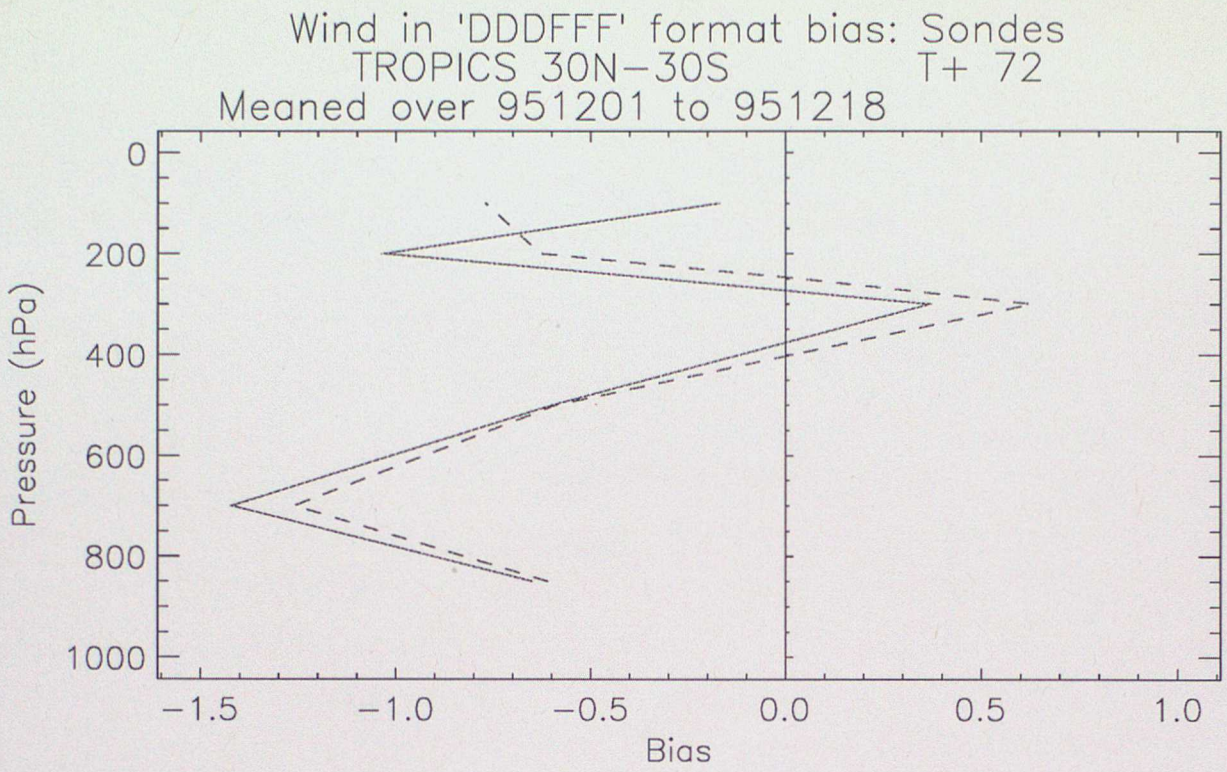


Fig. 5 Area 2 (LAM domain) pmsl verification against observations (lines as Fig. 1)



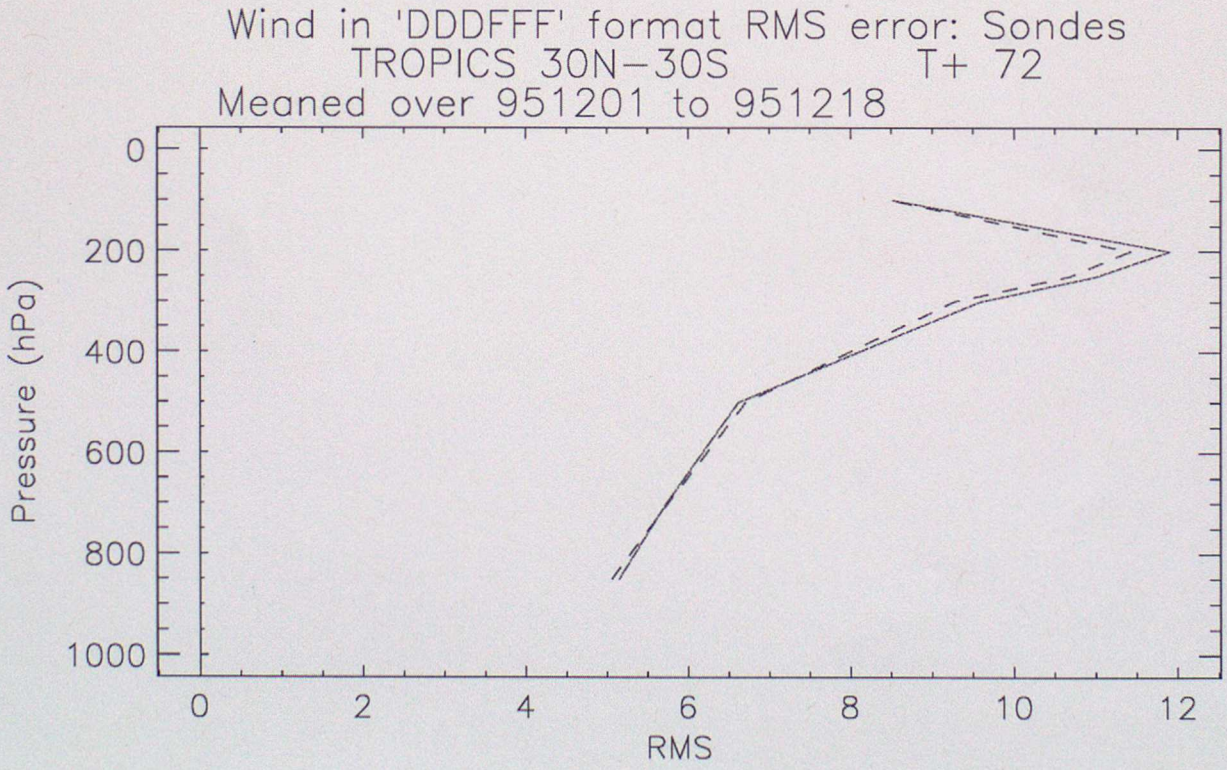
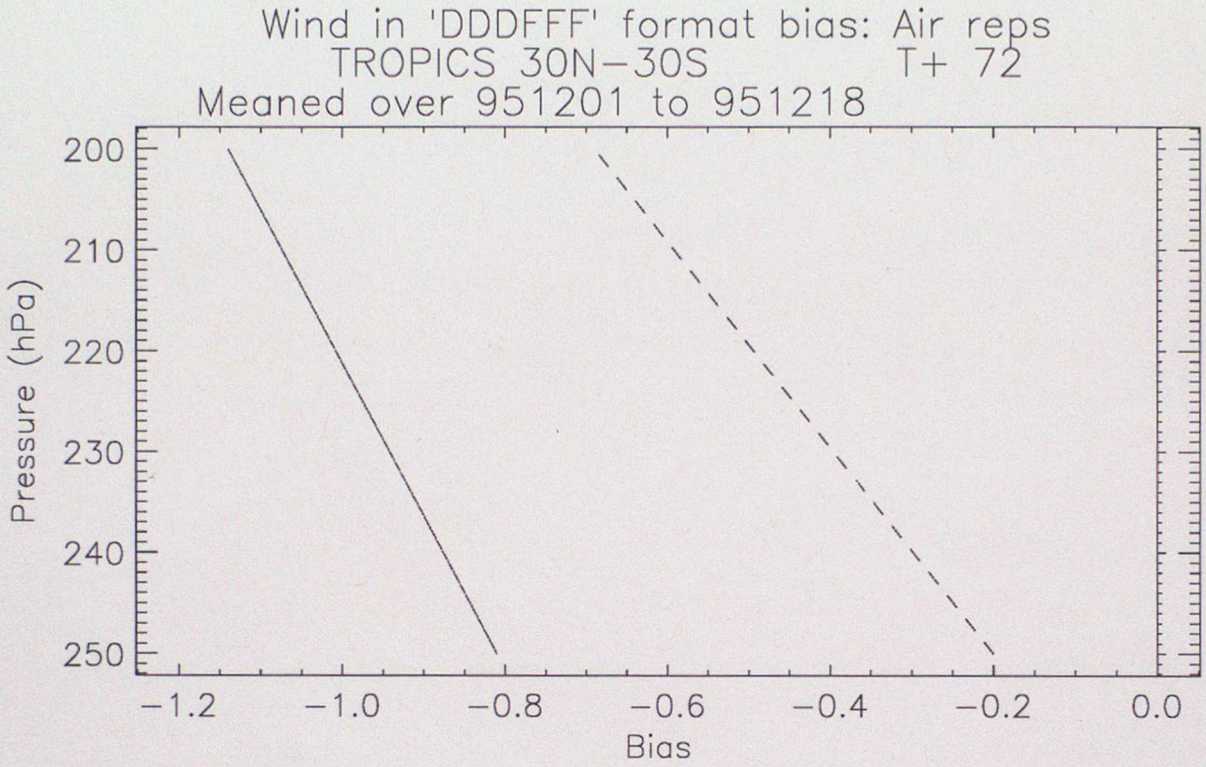


Fig.6 Windspeed verification against radiosonde for tropics at T+72



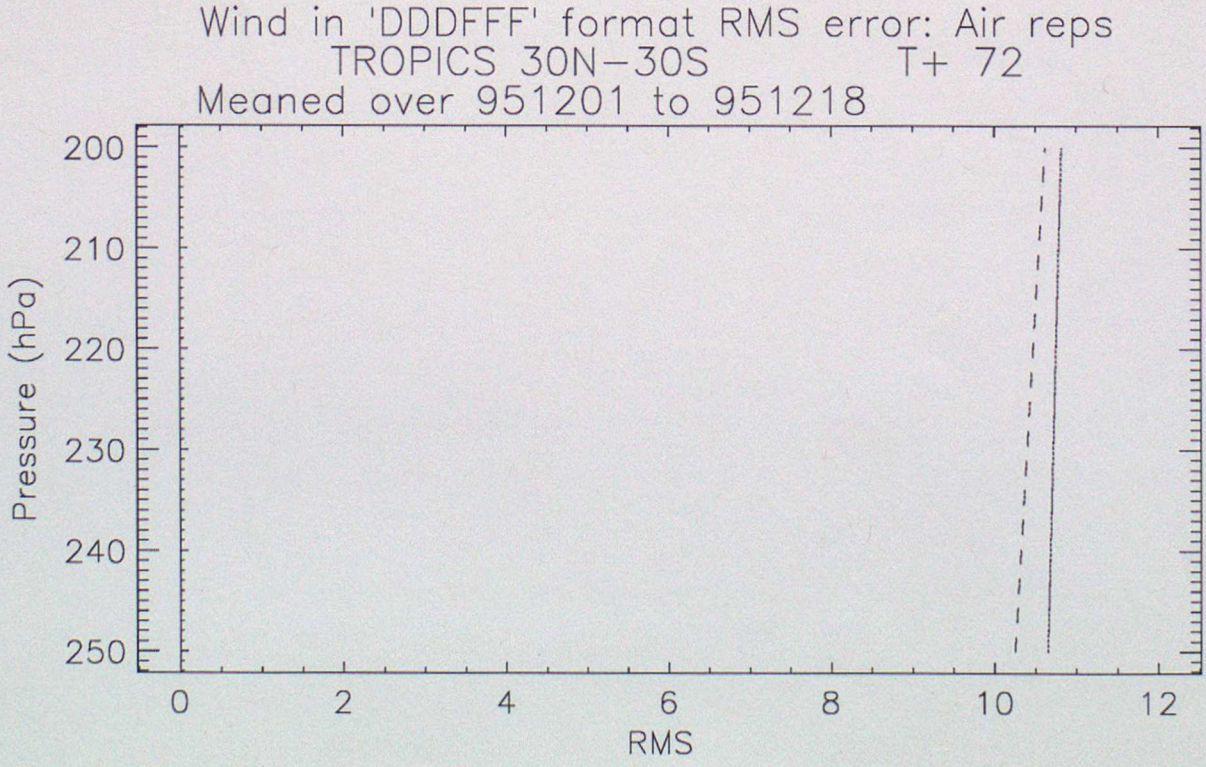
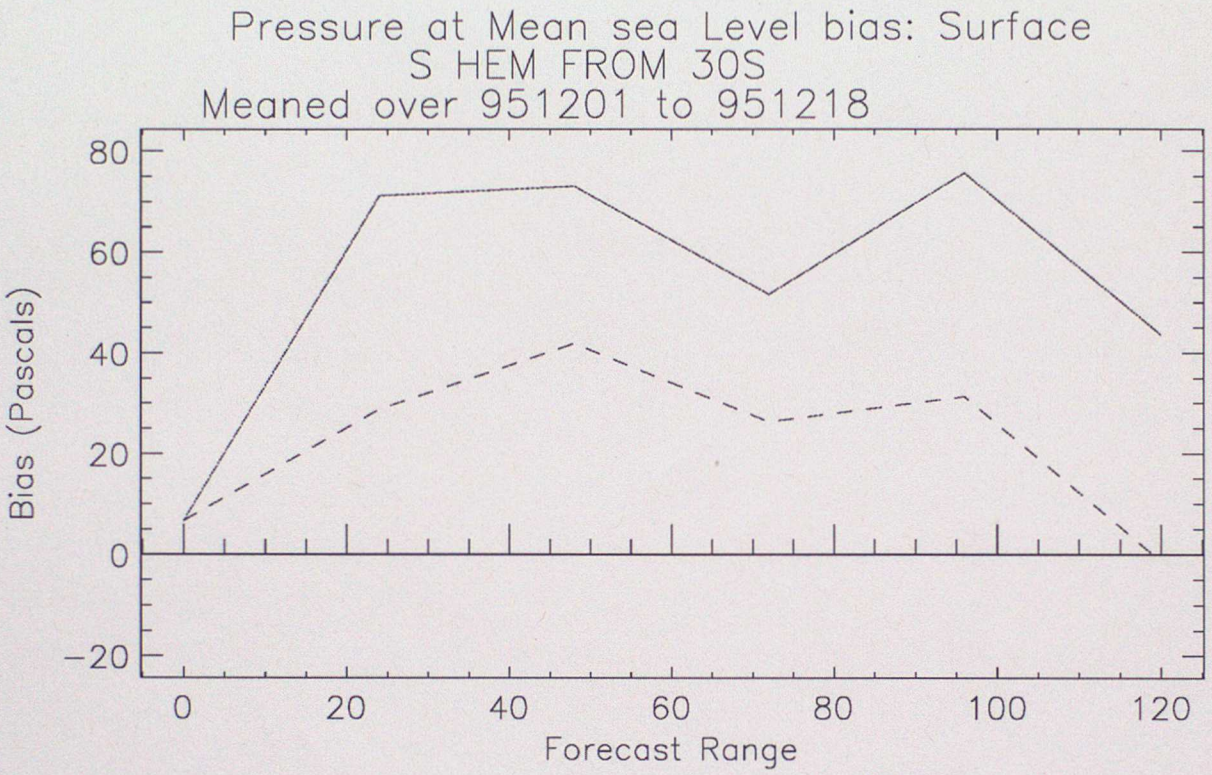


Fig.7 Windspeed verification against aireps for tropics at T+72



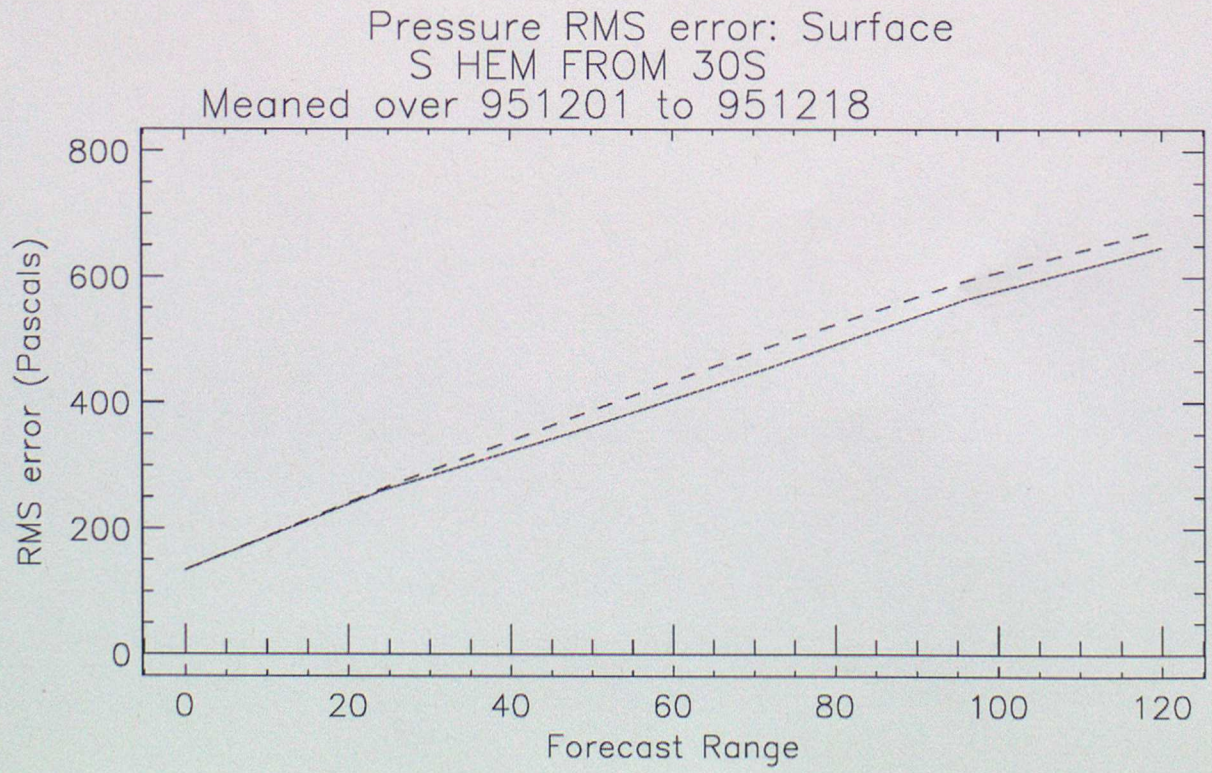
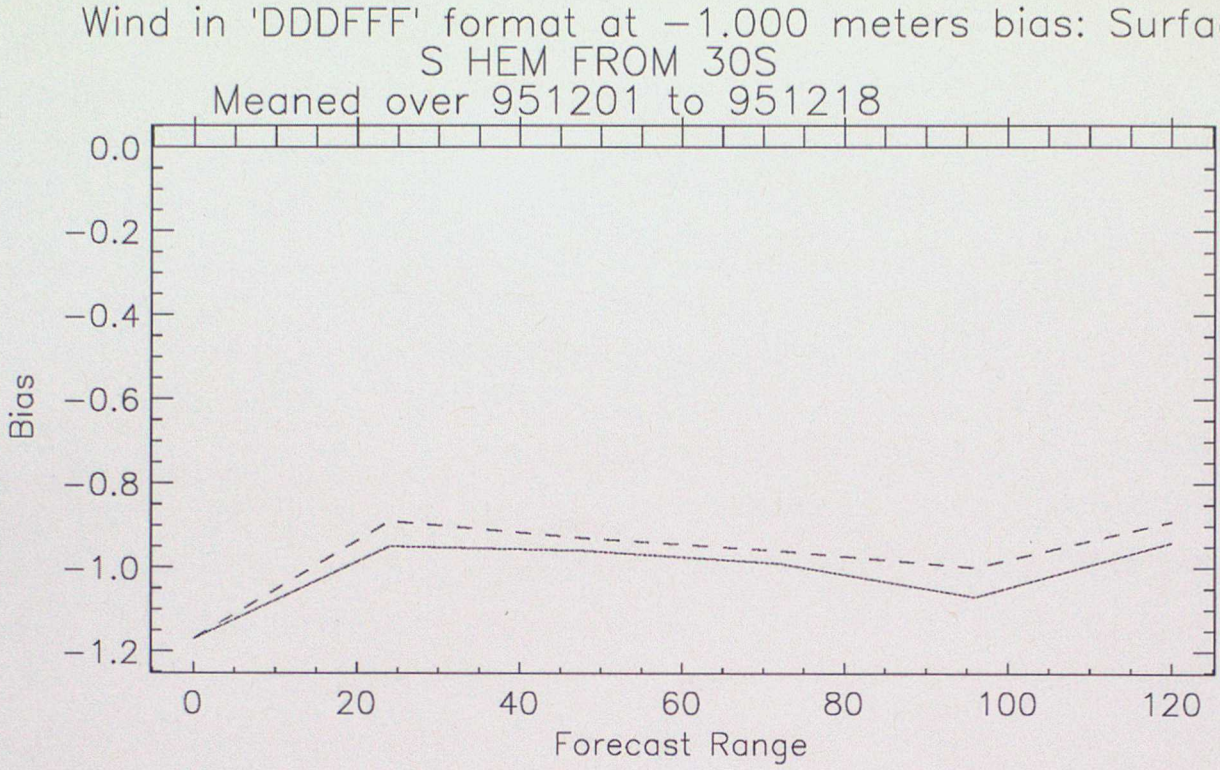


Fig. 8: Verification of pmsl for SH against observations



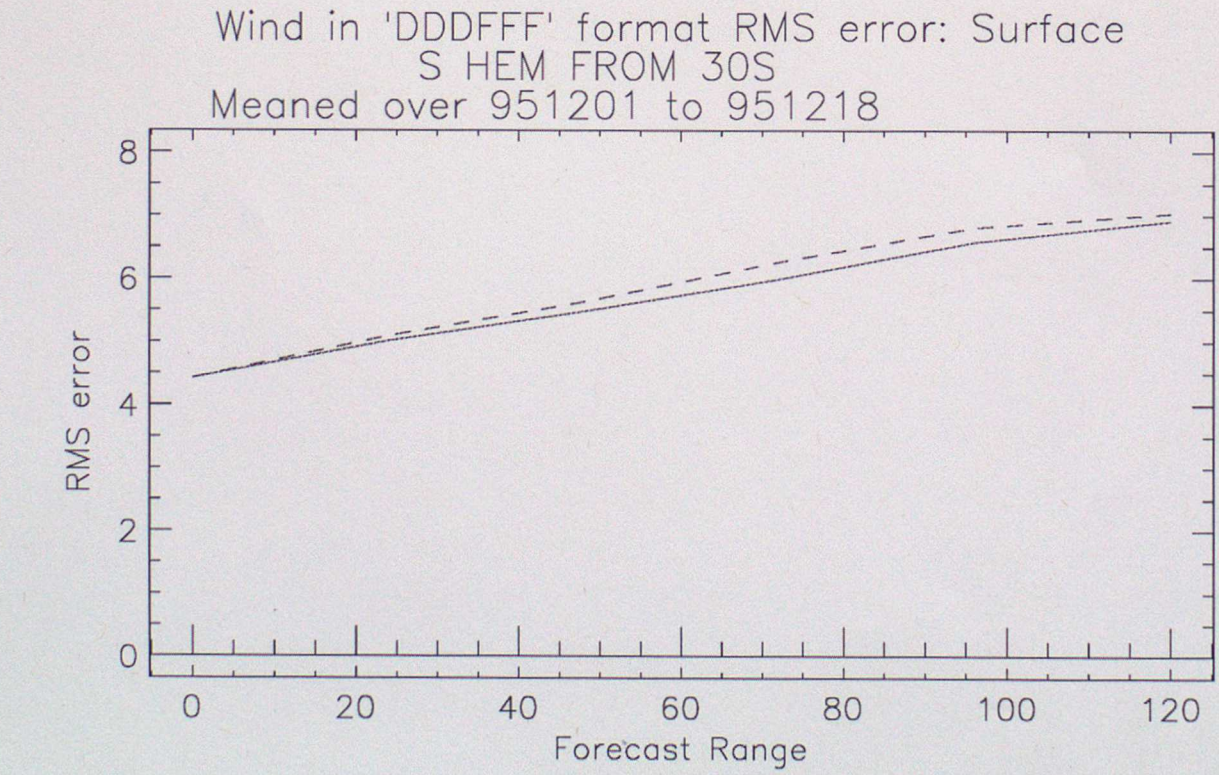


Fig. 9 Verification of surface wind for SH against observations

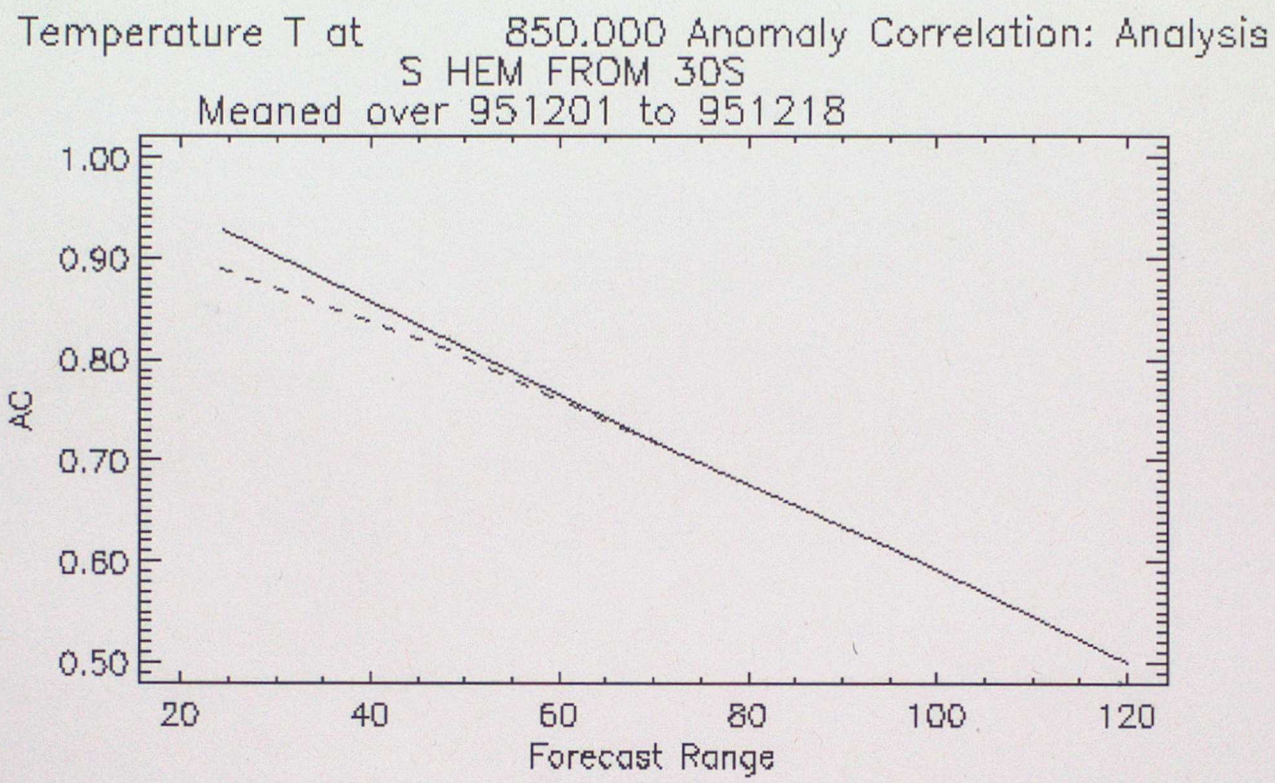
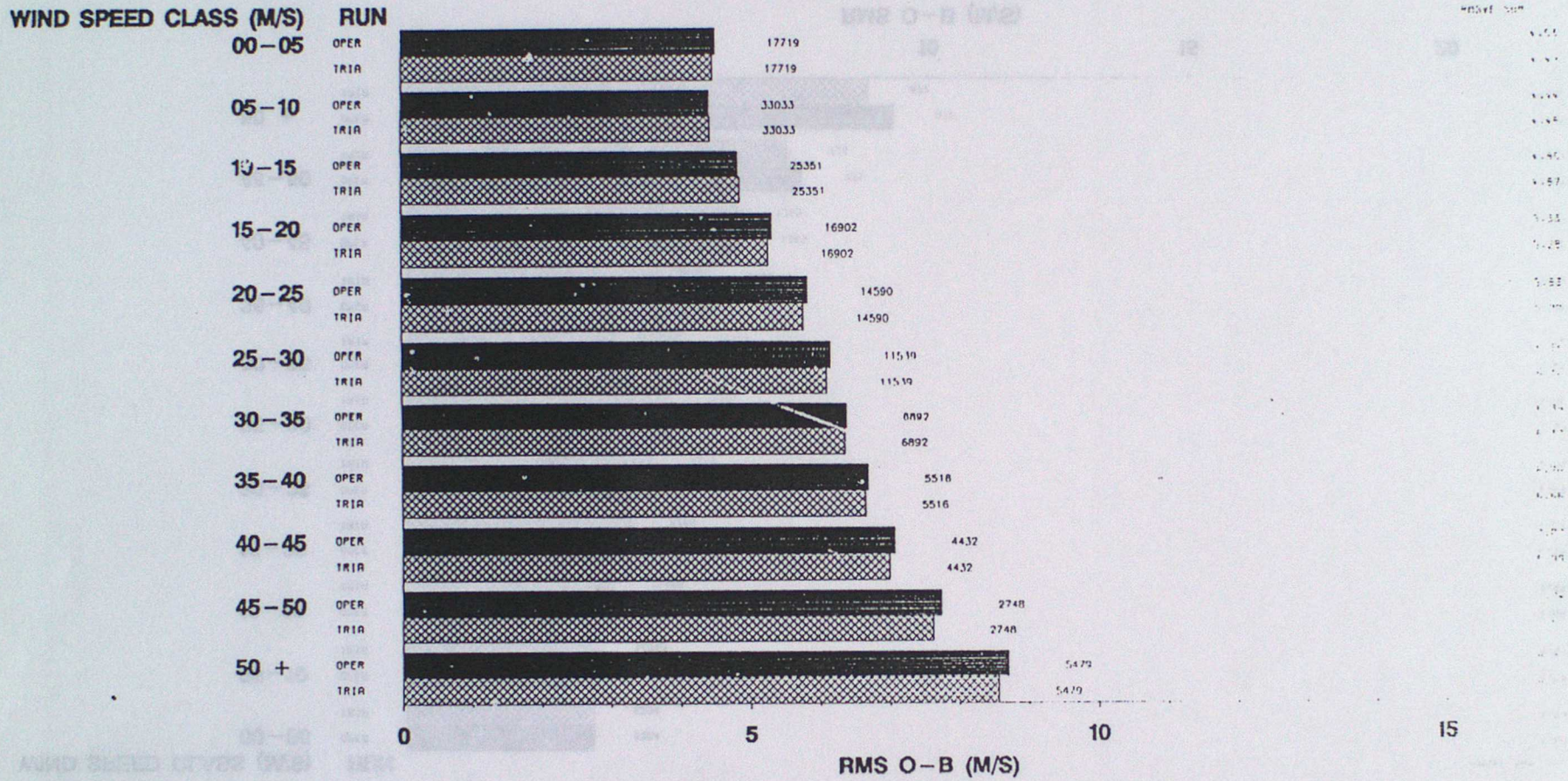


Fig. 10 SH AC for 850 hPa height

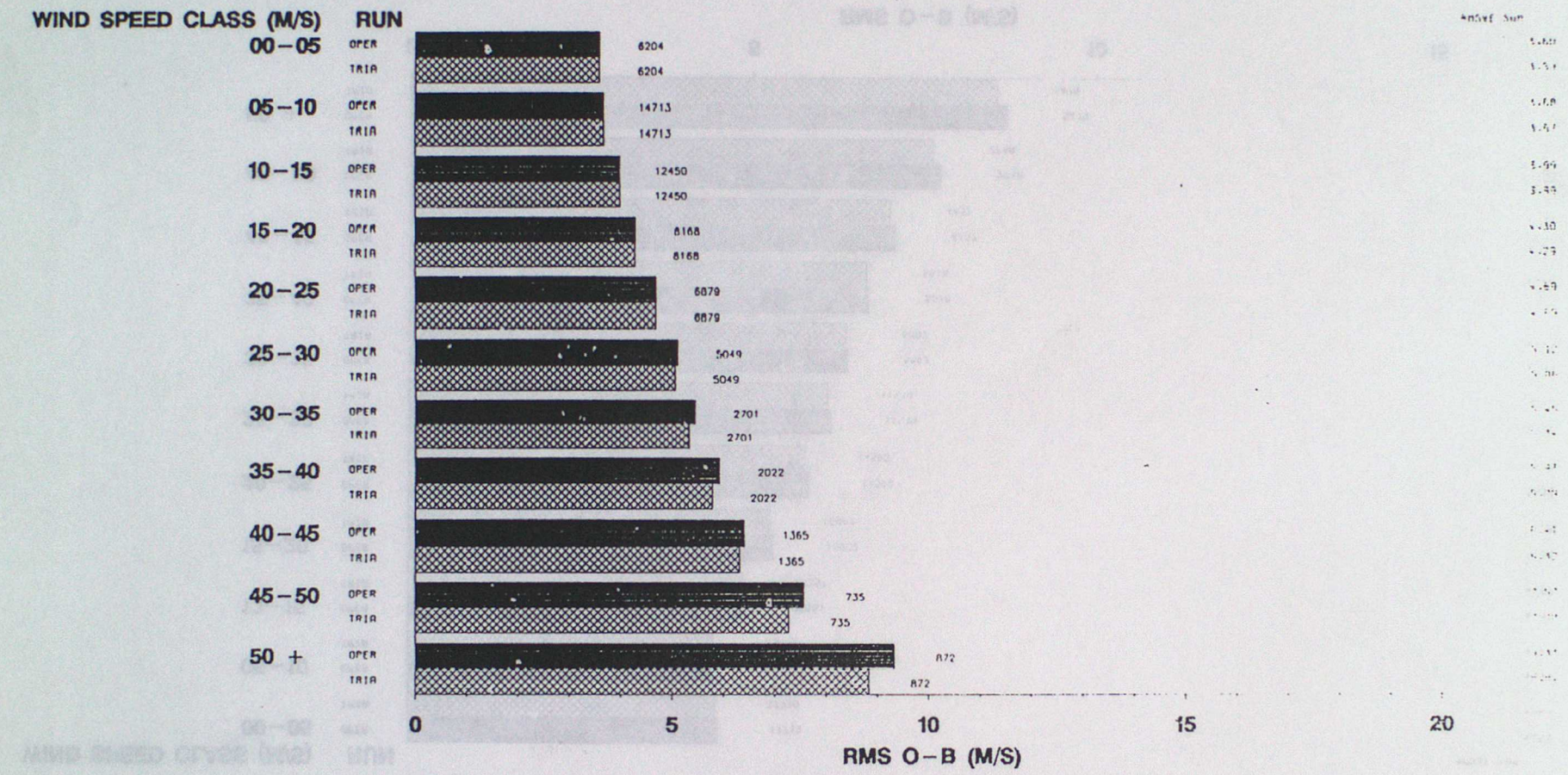
RMS VECTOR WIND DIFFERENCES (O−B) FOR OPER. AND TRIAL RADIOSONDES: GLOBAL



OBSERVATION SAMPLES PASSED QC IN BOTH RUNS

Annex 1

RMS VECTOR WIND DIFFERENCES (O−B) FOR OPER. AND TRIAL RADIOSONDES: NORTHERN HEMISPHERE 45−90

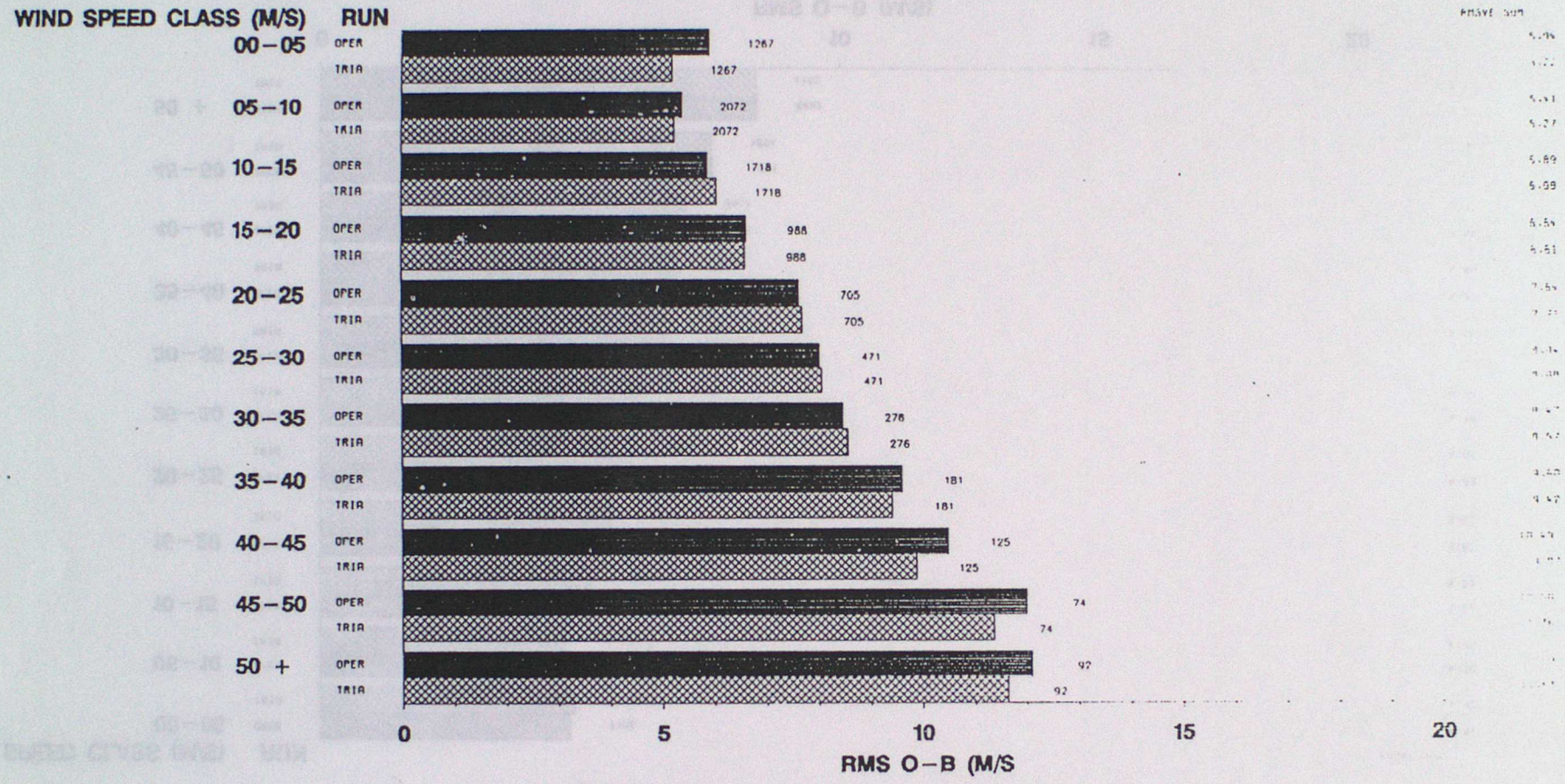


1−18 DEC.

PERIOD: 09-21 FEB 94
OBSERVATION SAMPLE: PASSED QC IN BOTH RUNS

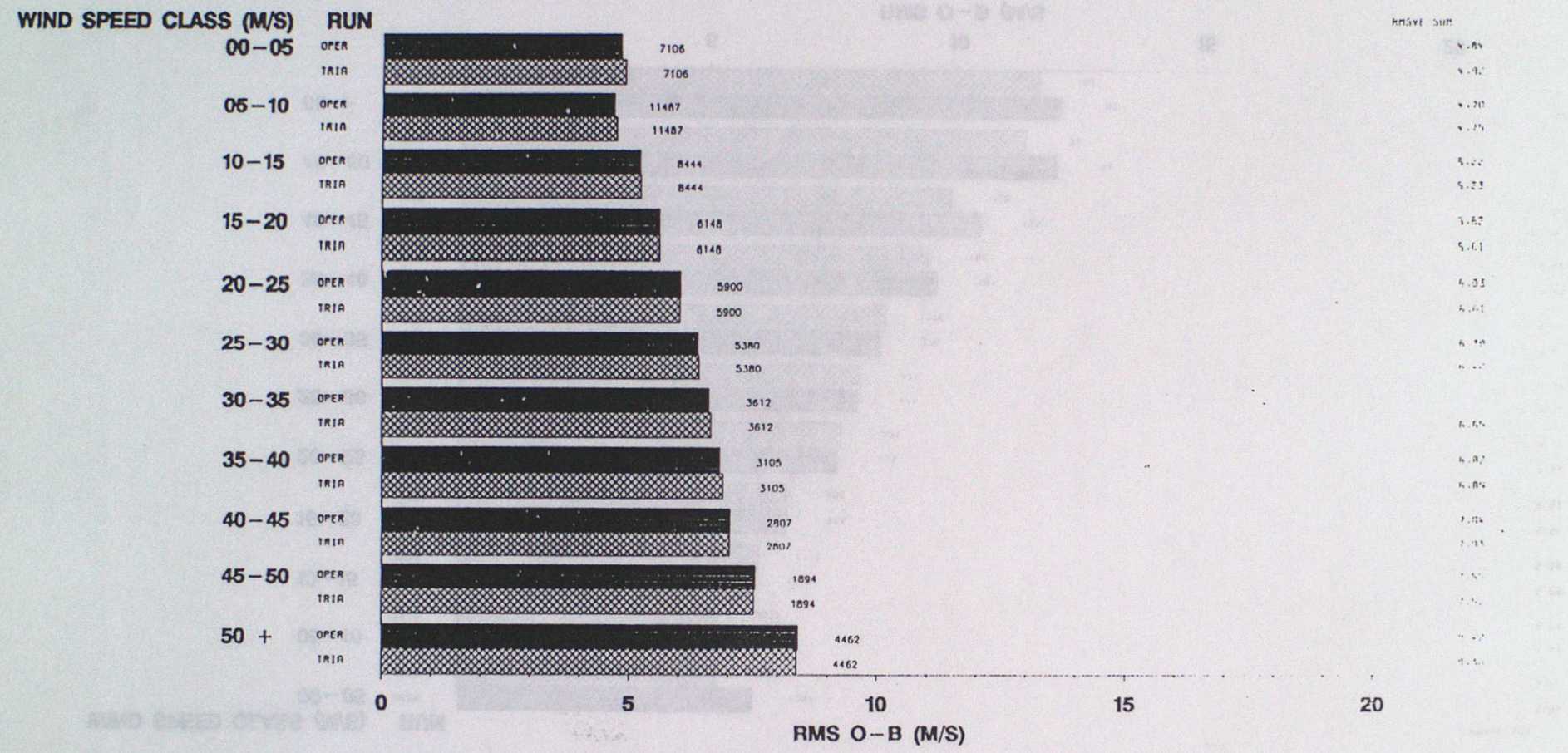
NB No 12Z data used. (not available for trial)

RMS VECTOR WIND DIFFERENCES (O−B) FOR OPER. AND TRIAL RADIOSONDES: SOUTHERN SUB−TROPICS (15S−45S)



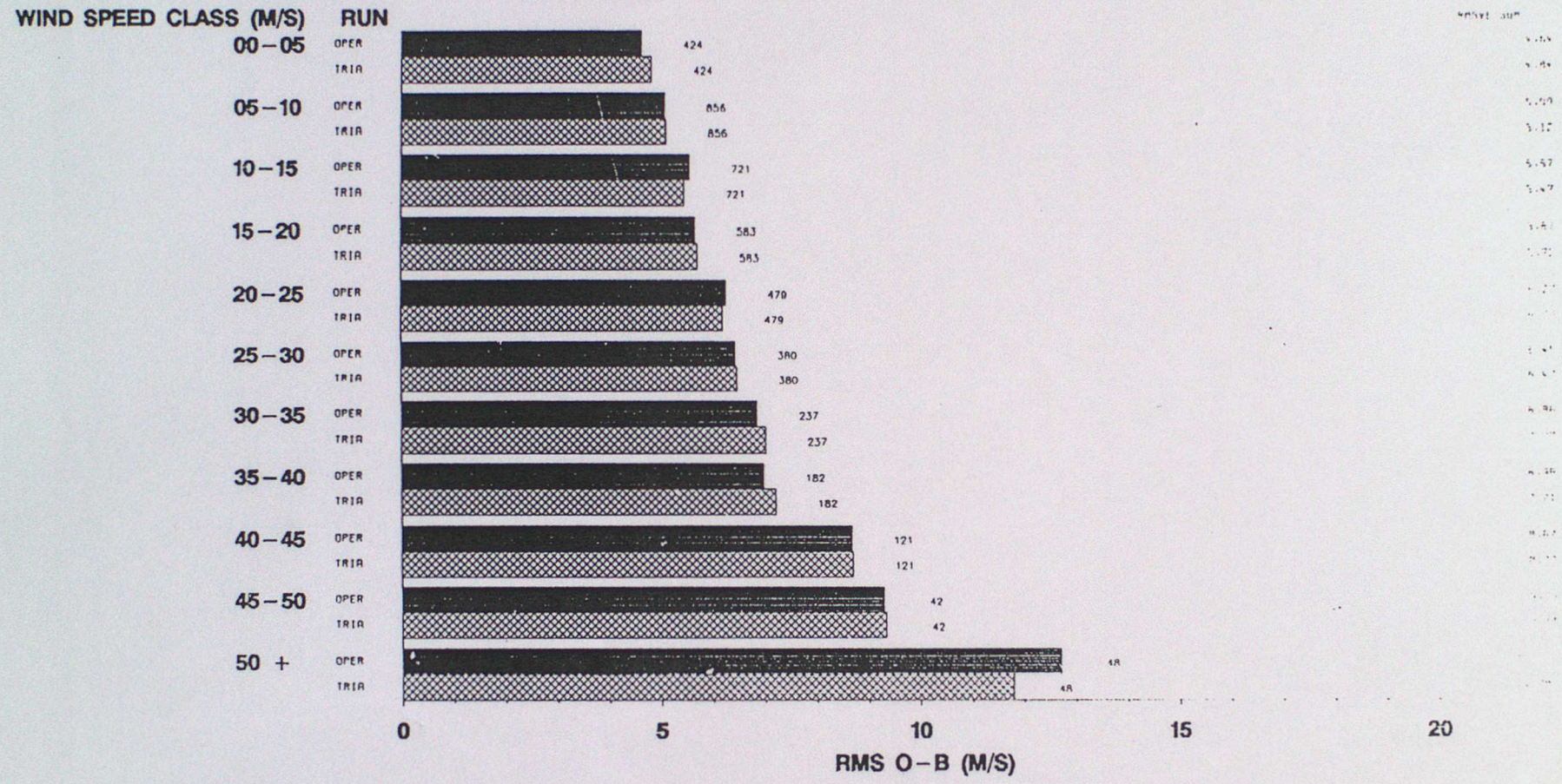
OBSERVATION SAMPLE: PASSED QC IN BOTH RUNS

RMS VECTOR WIND DIFFERENCES (O−B) FOR OPER. AND TRIAL RADIOSONDES: NORTHERN SUB−TROPICS (15N−45N)



OBSERVATION SAMPLES PASSED QC IN BOTH RUNS

RMS VECTOR WIND DIFFERENCES (O–B) FOR OPER. AND TRIAL RADIOSONDES: SOUTHERN HEMISPHERE 45–90



~~PERIOD: 09-04 FEB 94~~
OBSERVATION SAMPLE: PASSED QC IN BOTH RUNS

ANNEX 2

GLOSS Trial December 1995

Synoptic Assessment by J R Grant

Method

Forecasts at T+72, T+96 and T+120 were studied - these times in the forecast were chosen so that small differences at T+0 would have time to affect the evolution. The output was charts of GLOSS forecast, Operational forecast (Op) and the difference between the two. The forecasts for the southern and northern hemispheres were compared with the GLOSS and Op analyses by eye for MSL pressure and 250 mb heights. The Tropics were not assessed. Approximately 3/4 of the output has been examined. Selected cases are given to illustrate the main points.

The following bulletted points are conclusions with comments below.

- Overall GLOSS forecasts seemed to be as good as the Op.
  However, the details are mixed: at some times and in some places GLOSS was better, at other times and places GLOSS was worse than Op.

- Subjective judgment showed no significant improvement of the GLOSS forecasts over the Op.
  Occasionally there was a major error in both the forecasts, for example at T+120 an error of 45 mb in MSL pressure over Russia (Fig. 1). On this occasion GLOSS was slightly less bad. On another occasion, a south Atlantic low development was incorrectly forecast in both (VT 8/12). I have not the details of what GLOSS information was included in the assimilation for these forecasts - there may have been no up-to-date GLOSS just prior to the forecast run in the relevant area. An example of a better GLOSS forecast is shown in Fig. 2 with two troughs near Newfoundland and the build of high contours near UK at 250 mb T+96 valid 13/12.

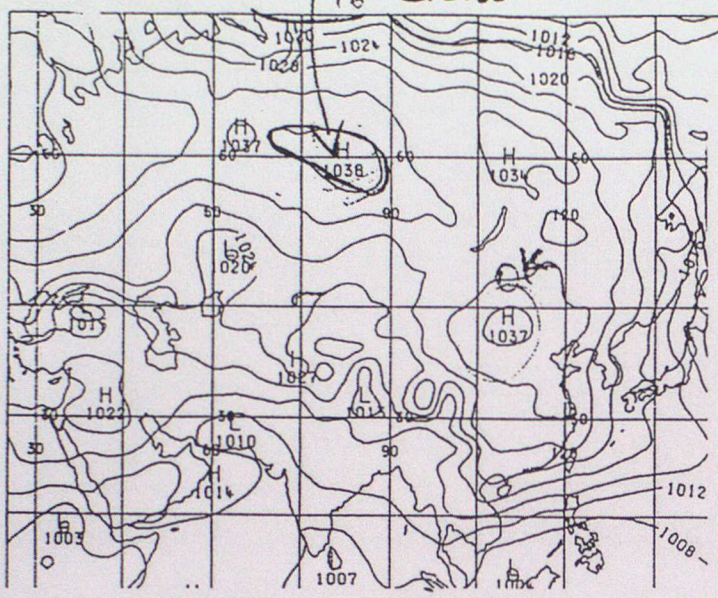
- On most occasions GLOSS seemed to be able to keep some of the smaller scale detail better than in Op.
  Note that this may not always reduce the rms error as a slight misplacement of the feature, especially in a strong airflow, might increase the rms error. An example of better shape to a trough in GLOSS occurred in the S Atlantic at 250 mb T+96 valid 11/12 (Fig. 3).

- It is surprising that, in the southern hemisphere especially that, in general, the T+0 charts showed little difference between the GLOSS and Operational analyses.
  (There were, however, occasions when differences of -8/+9 occurred at T+0 - 250 mb S Pacific on 11/12). This is presumably because the Op receives the SAT 120 or 500 data. It raises the question of how GLOSS is assimilated - is each reading spread out 400 km or so in each direction? If so, this means that the detail that the GLOSS data could provide may be being swamped by averaging. I am thinking for instance of dry zones associated with cyclogenesis which can be 100's of km long but only 10's of km wide.

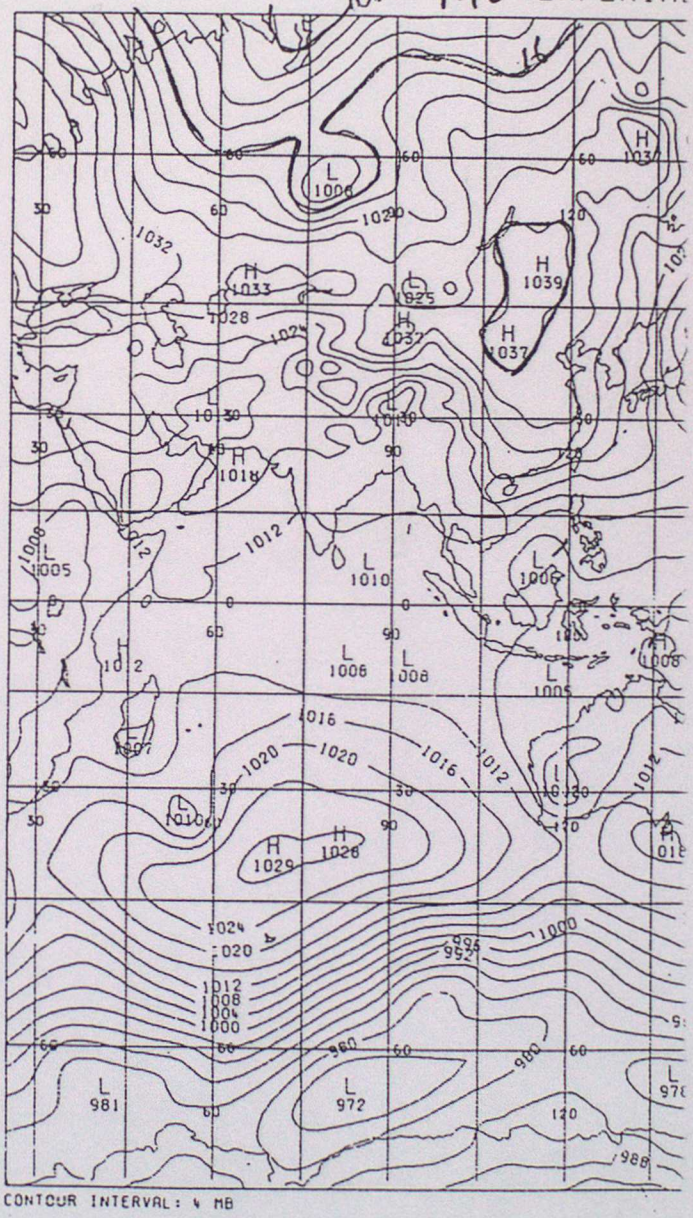
J R Grant
8 January 1996

Fig. 1

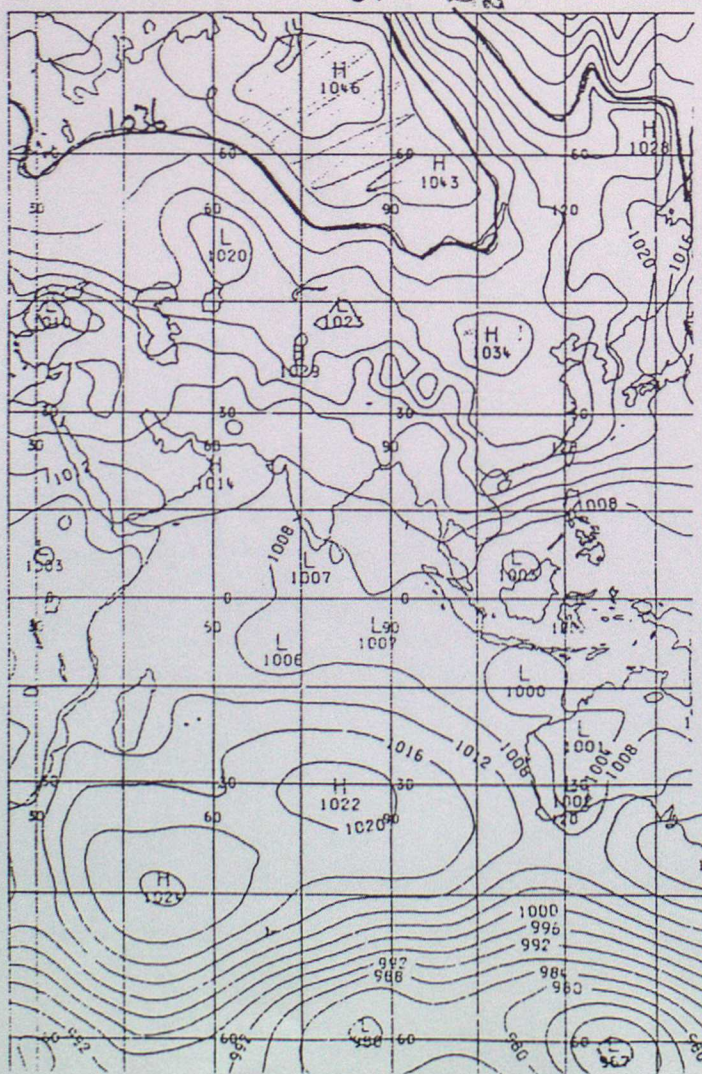
SIXTH GLOSS TRIAL: PARALLEL RUN T+120
MEAN SEA LEVEL PRESSURE (GLOSS TRIAL)
VALID AT 12Z ON 6/12/1995 DAY 340 DATA TIME 12Z ON 1/12/1995 DAY 335
SEA LEVEL GLOSS EXPERIMENT NO.: 1 T+120



SIXTH GLOSS TRIAL: PARALLEL RUN
MEAN SEA LEVEL PRESSURE (GLOSS
VALID AT 12Z ON 6/12/1995 DAY 3
SEA LEVEL T+0 EXPERIM



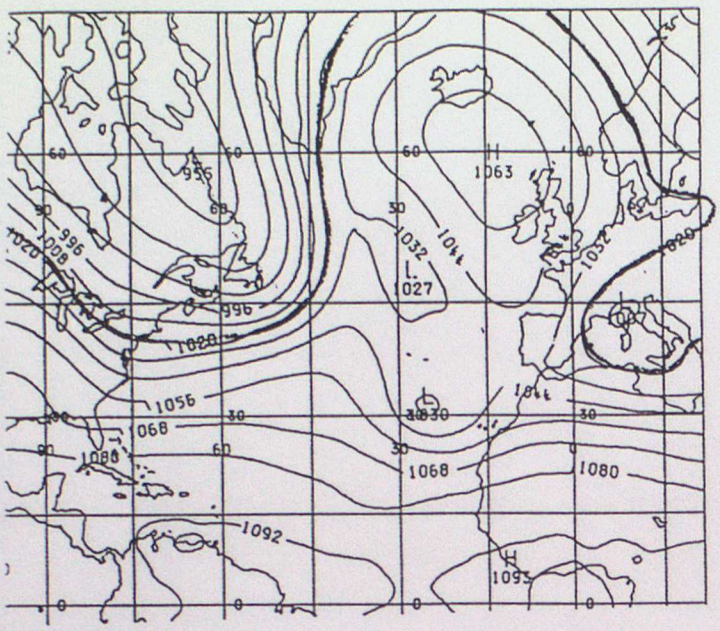
SIXTH GLOSS TRIAL: OPERATIONAL F
MEAN SEA LEVEL PRESSURE (CONTROL
VALID AT 12Z ON 6/12/1995 DAY 34
SEA LEVEL OP EXPERIME



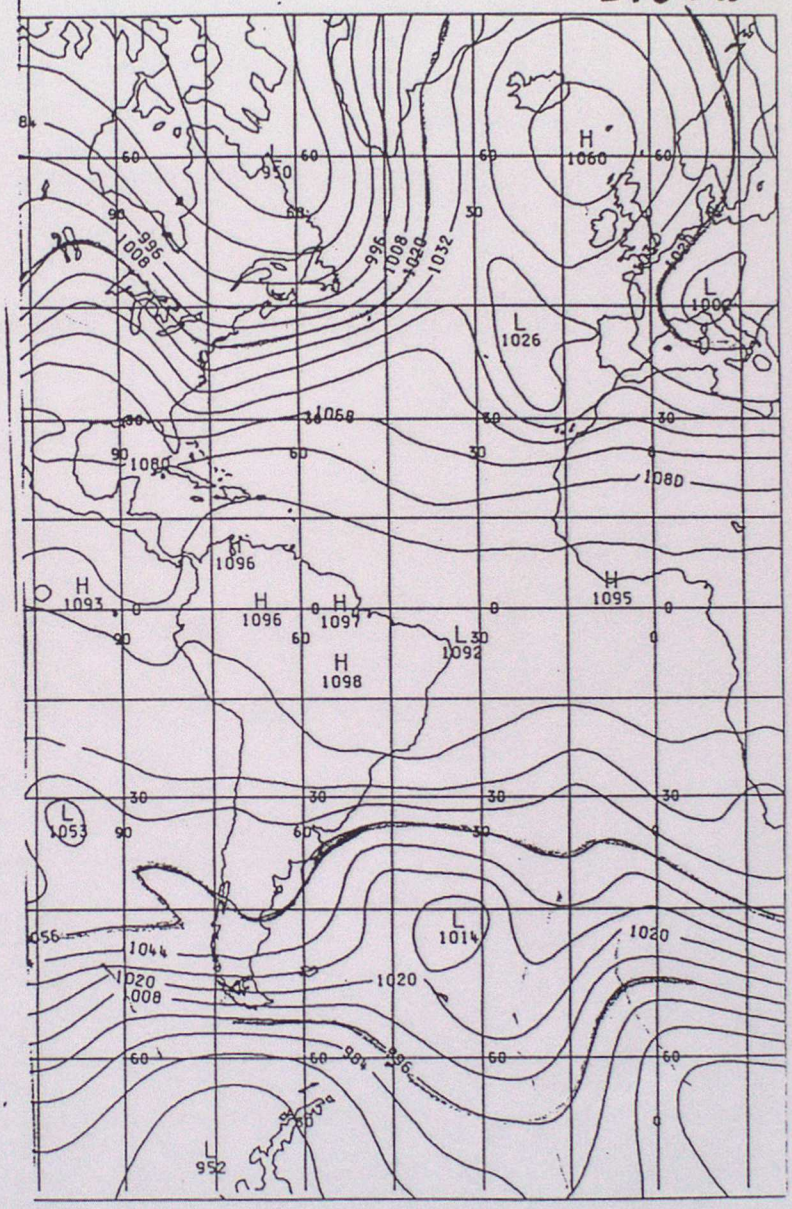
CONTOUR INTERVAL: 4 MB

Fig. 2.

343 Gloss T+96 250 mb



DAY 347 Gloss T+0 13/12/95 1200 250 mb



343 OP T+96 250 mb

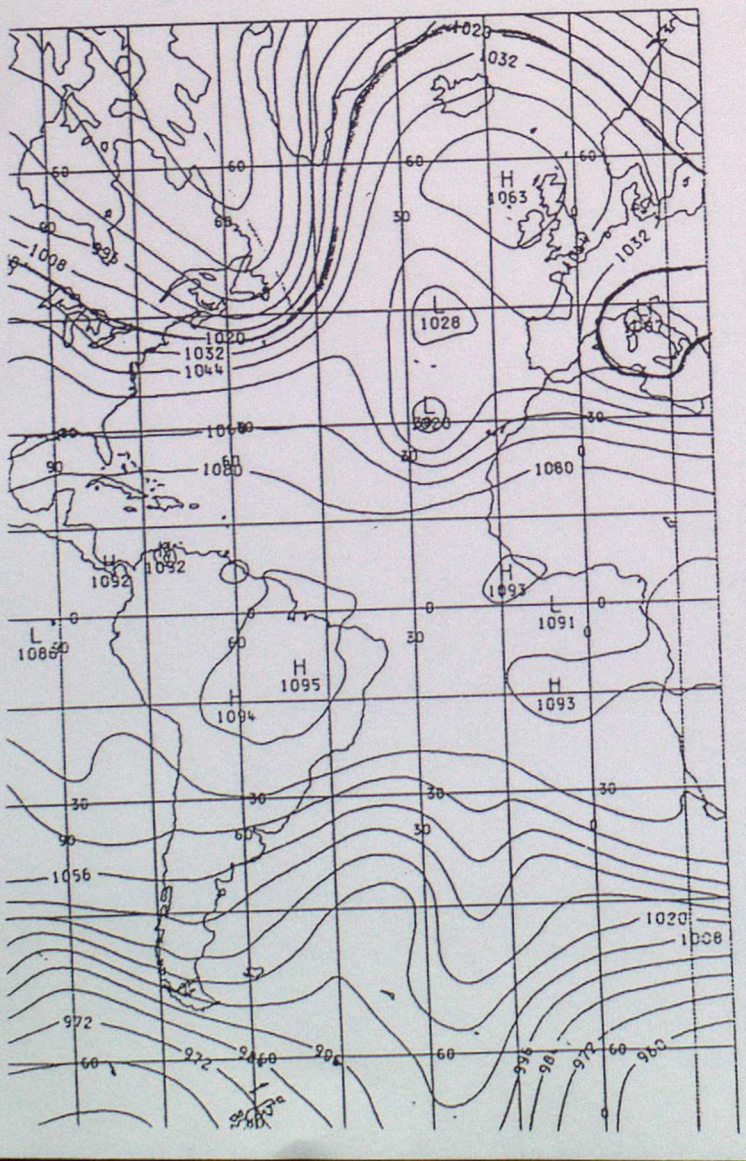
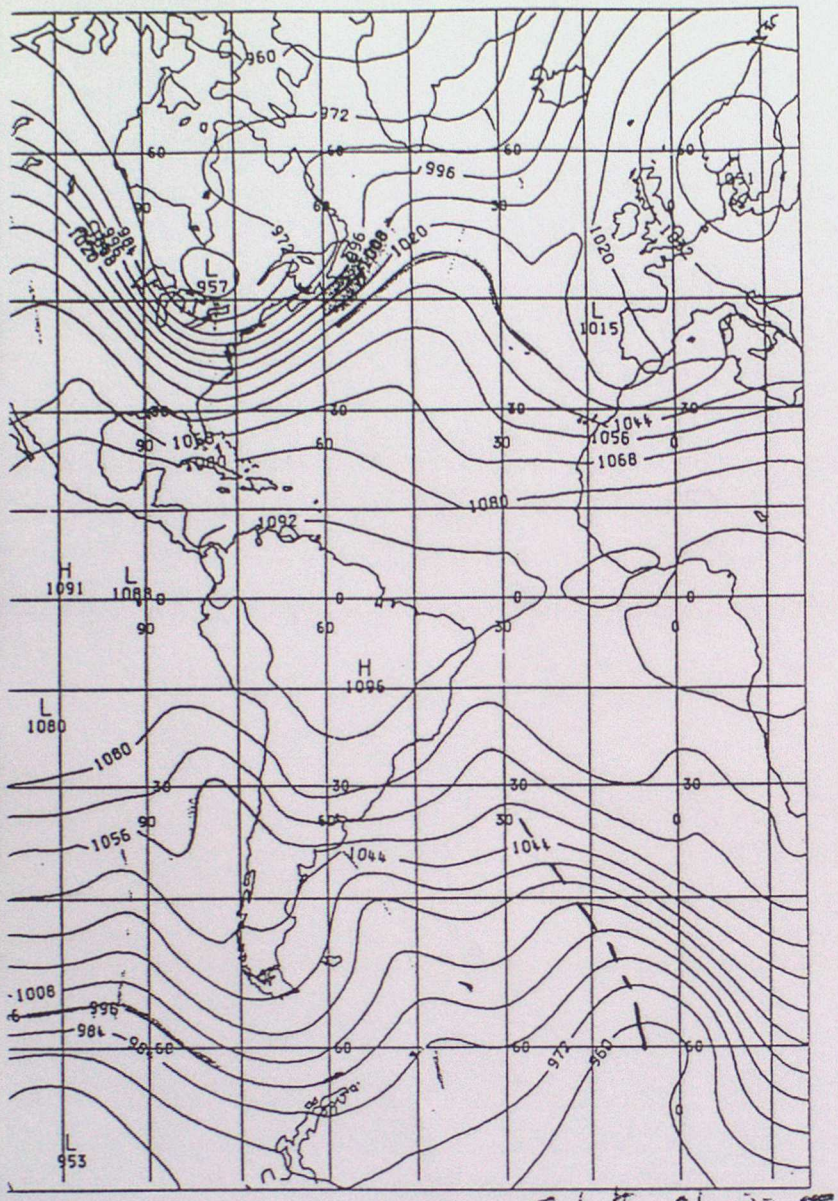


Fig. 3.

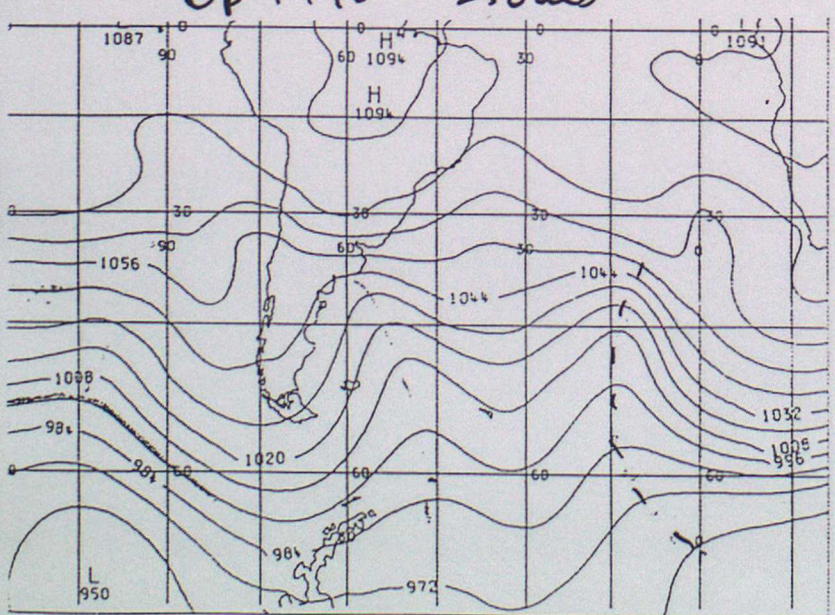
DAY 341

GLOSS T+96 250 mb



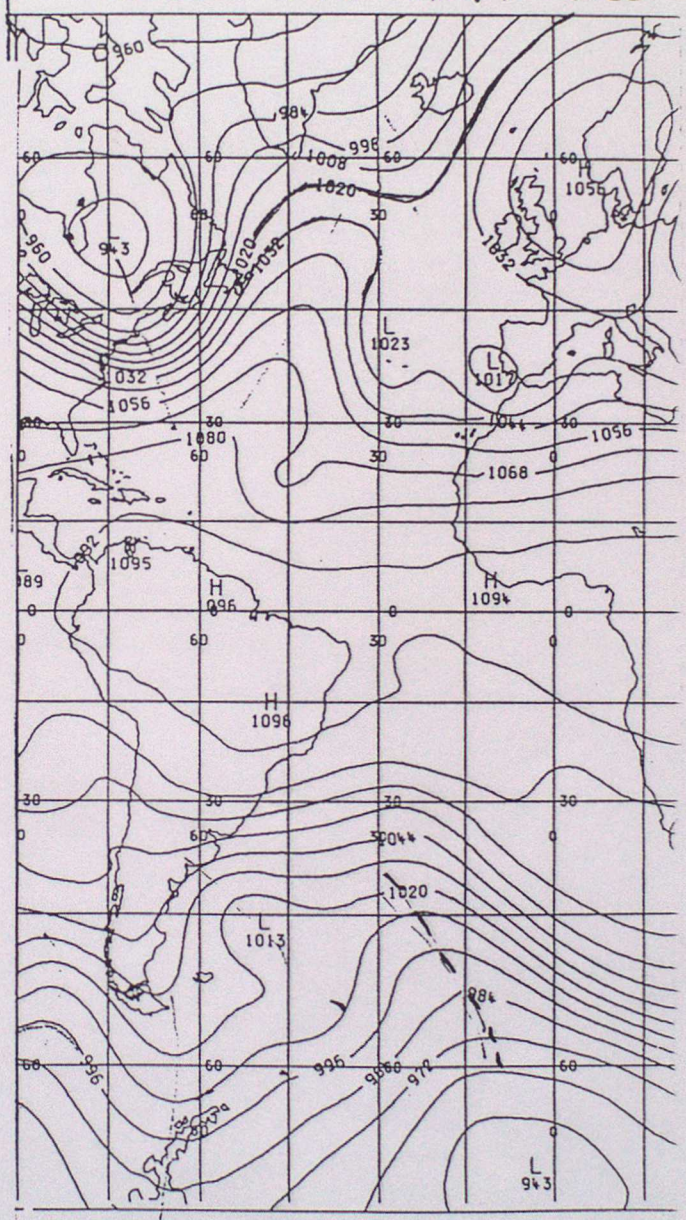
G better shape

Op T+96 250 mb



345

GLOSS T+0 11/12/95 1200



Annex 3

PARALLEL SUITE EXPERIMENT No.37

Tropical Cyclone Forecast Track Verification

1. Sponsor

Brian Barwell, NWP Division.

2. Description of Experiment

The parallel suite PG12 global model assimilation includes Gloss data (Gloss experiment #6).

3. Duration and Forecast Specification

The total period of the experiment was from 28.11.95 to 18.12.95
PS forecasts were run out to T+120 every 24 hours at 12Z.

4. Tropical Cyclone Forecast Verification

a) Differences in samples
The PS run DT 12Z 10.12.95 dissipated TC Frank by T+48; the corresponding OS run did not.

b) Verification Statistics
Total number of storms which were active during the period : **2**
Homogeneous samples were used in the statistical comparison

Forecast Error comparison for whole globe

	T+0	T+24	T+48	T+72	T+96	T+120
No. verified	9	7	4	2	2	1
Operational DPE	44	153	296	551	676	1038
Parallel Suite DPE	35	163	255	473	556	533
% increase	-20.5	+6.5	-13.9	-14.2	-17.8	-48.7

No. verified : The number of forecasts verified for each forecast time
DPE : Direct Positional Error in kilometres
% increase : Percentage increase in PS analysis/forecast error compared to OS

The overall percentage increase in PS analysis/forecast error : -12.3%

5. Comments

There were only two storms in the south-east Indian ocean during the trial, so the sample size for verification was very low. However, in the small number of forecast verified, the Gloss data appeared to have a significant positive impact at all times except at T+24. Of the 25 individual analysis/forecasts verified the PS was better than the OS in 17 and the OS better than the PS in 6. They were equal in the other two.

The large difference in the one T+120 forecast verified was at DT 12Z 07.12.95 for TC Frank. The OP forecast took the storm south towards the Australian coast up to T+72, then turned it back on itself at T+96 and T+120. The PS forecast took a more west of south track initially and then turned the storm south-westwards at later forecast times. This was much closer to reality.

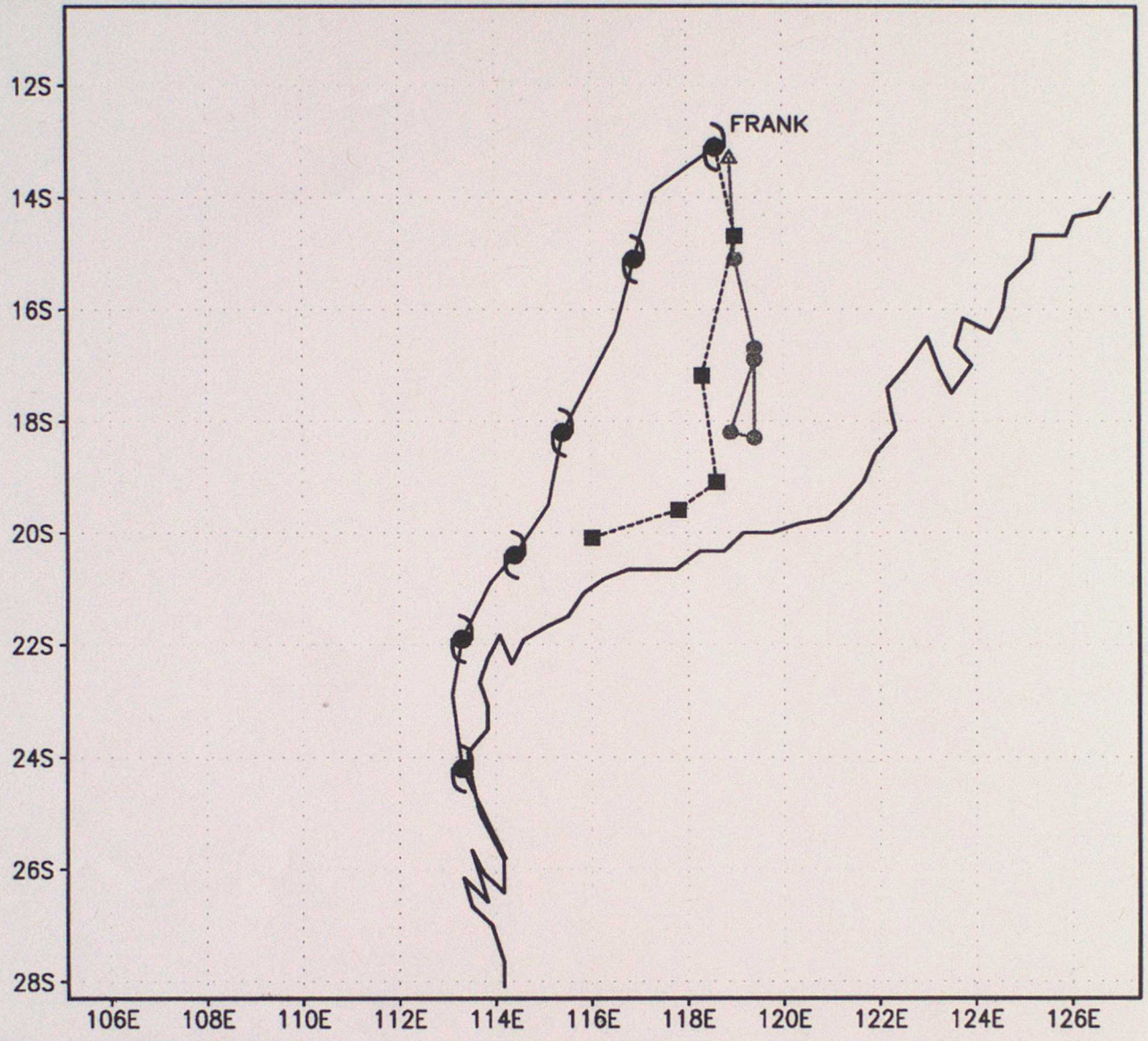
Although the small number of cases reduces the statistical validity of the results, Gloss data appears to have had more than just a minor impact on the forecast tracks of these tropical cyclones.

Details of the formulation of the UKMO TC verification Scheme can be found in NWP Gazette Vol.1, No.2. December 1994.

Julian Heming, FS9

UKMO GLOBAL MODEL and TEST

FORECAST TRACKS of TROPICAL CYCLONE FRANK



UK 951207 TE 951207

KEY to FORECAST TRACKS

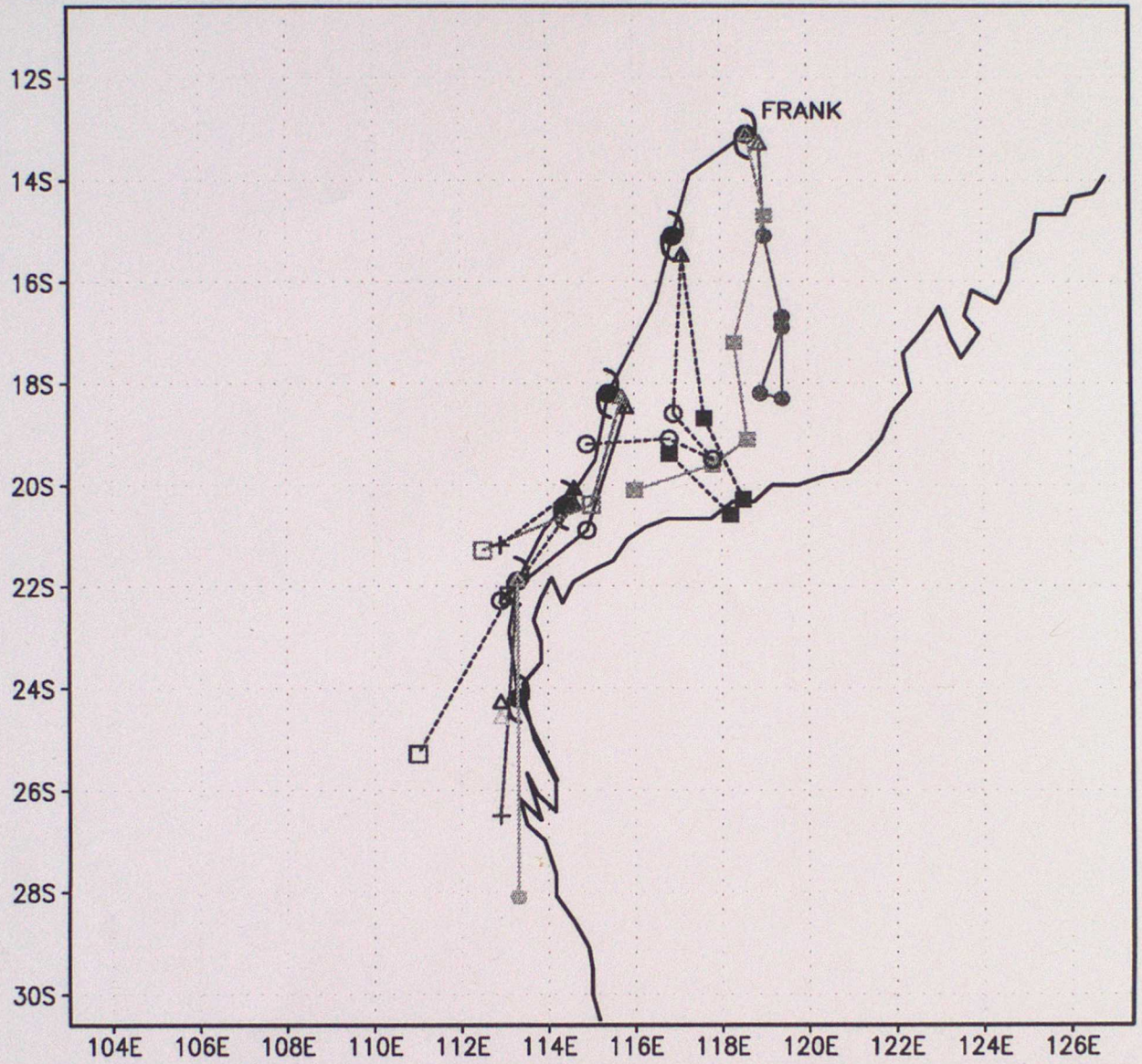
(Triangles denote analysed positions)

24 HOURLY OBSERVED POSITIONS

DATE/TIME OF FIRST SYMBOL 12Z 07 DECEMBER 1995

UKMO GLOBAL MODEL and TEST

FORECAST TRACKS of TROPICAL CYCLONE FRANK



UK 951207 UK 951208 UK 951209 UK 951210 UK 951211 UK 951212

TE 951207 TE 951208 TE 951209 TE 951210 TE 951211 TE 951212

KEY to FORECAST TRACKS

(Triangles denote analysed positions)

24 HOURLY OBSERVED POSITIONS

DATE/TIME OF FIRST SYMBOL 12Z 07 DECEMBER 1995

Annex 4

GLOSS BIASES IN THE ARCTIC IN THE NORTHERN WINTER

B R Barwell

When GLOSS was first assimilated operationally in January 1994, increases in tropospheric temperature were observed, especially over Arctic Canada. As these were judged to be unrealistic by forecasters, GLOSS was removed after about one week. Since tropospheric HIRS are not used for retrievals over sea ice, the problem could be traced MSU-2 since this channel was the dominant one for determining the tropospheric retrieval.

As an illustration of the problem, Fig. 4.1 shows charts of the mean difference between observation and background for NOAA-12 MSU-2 for a 3-week period in January 1994. Figure 4.1a shows the mean (O-B) difference chart before bias correction and 4.1b is the corresponding chart for bias corrected observations (C-B). Values in degrees K are given for 15-degree boxes on each chart. These charts were produced in early February 1994 and used the version of GLOSS and the bias correction procedure operational at that time. The uncorrected observations show maximum differences of about 3.2K over the Canadian Arctic and 3.7K over eastern Russia. In the tropics there is a small bias of about 0.5K. Figure 4.1b shows that, while the tropical bias is largely removed by the bias correction, the large differences in the northern hemisphere increase to 3.9 and 4.4K respectively even though the purpose of the bias correction is to reduce the mean differences from background.

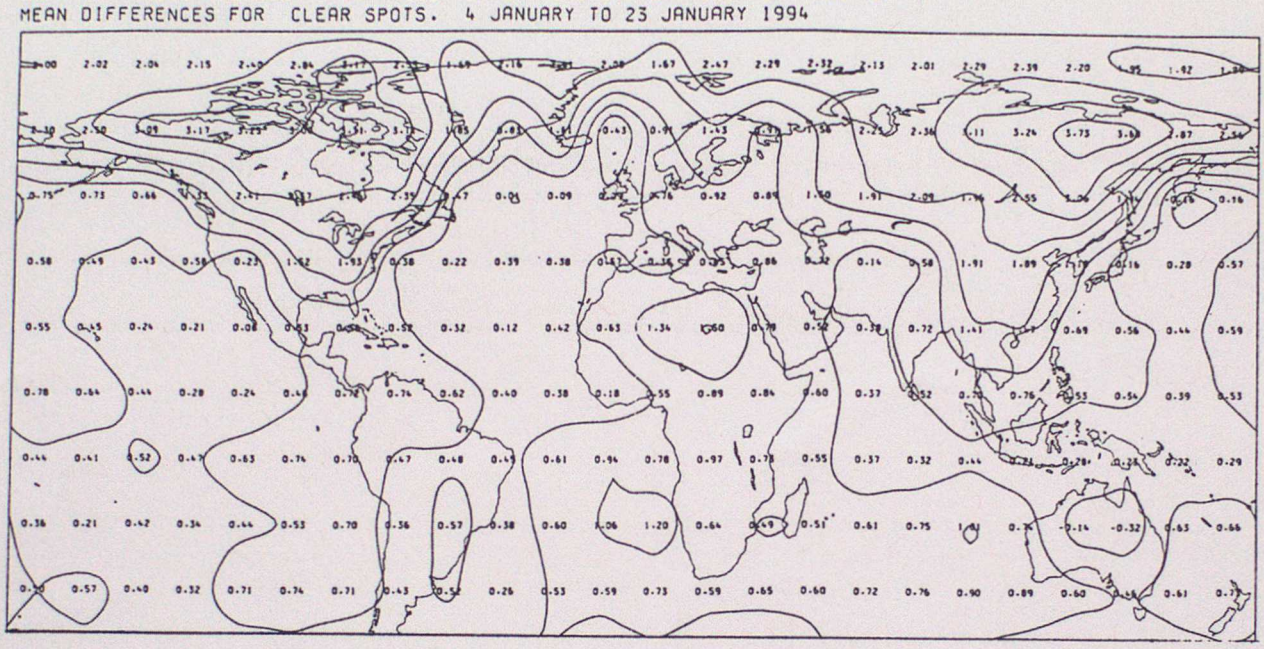
Figure 4.2 shows the situation in January 1996 using the version of GLOSS used in the parallel trial. Figs. 4.2a and 4.2b show maps of (O-B) and (C-B) corresponding to Fig. 4.1, and Fig. 4.2c shows the mean bias corrections applied. Data is again for the MSU-2 channel on NOAA-12 and covers a period of about three weeks. Once again, (O-B) maxima are found over Arctic Canada and Russia but values are smaller than before, typically 2.5K and 3.7K respectively and the gradient of the contours round these two areas is much reduced from 1994 levels (Fig. 4.1). Contributing factors to this change are the scan-dependent bias correction (which could not be applied to 1994 data) and the recent modifications made to the background profile below 950hPa over sea ice designed to reduce (O-B) differences mainly through removal of a suspected bias in the model's surface temperature.

In addition to improvements in (O-B), Fig. 4.2b shows that the new bias correction coefficients (which are now computed using data over sea ice as well as sea) make further reductions in difference from background in contrast to the increases found in 1994. In the two main areas, levels are reduced to around 1.6K and 2.2K respectively while the smaller bias in the tropics is removed as before. Figure 4.2c confirms that the bias corrections are all negative this time and that the magnitude increases in the polar regions instead of changing to positive values. These results are confirmed by data from the rerun of GLOSS for the period 25 January to 3 February 1994 (not shown). In the region of Hudson's Bay in the control run (for which the original bias correction coefficients were used) the correction increases the difference from background for NOAA-12 MSU-2 from 2.25k to nearly 4K. In the GLOSS run with newly computed coefficients there is a decrease in bias of about 0.7K throughout the north polar regions.

Clearly,even with GLOSS in its present form, some warming of the troposphere in the polar regions is required to fit the MSU-2 observations. It is not possible to tell whether the cause is some data bias due to the NESDIS emissivity correction for example, or whether it is the result of a model bias such as excessive cooling at or near the surface, or some combination of the two effects. Evidence from the "30-level + GLOSS" trial suggests that warming occurs close to the surface and spreads upwards through the troposphere during forecasts. This may be aggravated by assimilation of the GLOSS

retrievals at 1000hPa where the background profile may be considerably warmer than the model forecast because of the adjustment made to surface and 1000hPa temperatures over sea ice. This can lead to large temperature increments being used in the assimilation at this level and thereby possible excessive heating. It is therefore recommended that over sea ice, GLOSS retrievals should not be assimilated below 950hPa. This can be achieved by setting the necessary flags within the GLOSS executable. To check that this change does not produce additional undesirable effects, the January 1994 rerun was repeated without using GLOSS at low levels over sea ice and was found to give only very minor changes to verification against analyses.

(a) NOAA-12 (OBSERVED-BKG) BRIGHTNESS TEMPERATURE DIFFERENCES FOR MSU 2
MEAN DIFFERENCES FOR CLEAR SPOTS. 4 JANUARY TO 23 JANUARY 1994



(b) NOAA-12 (CORRECTED-BKG) BRIGHTNESS TEMPERATURE DIFFERENCES FOR MSU 2
MEAN DIFFERENCES FOR CLEAR SPOTS. 4 JANUARY TO 23 JANUARY 1994

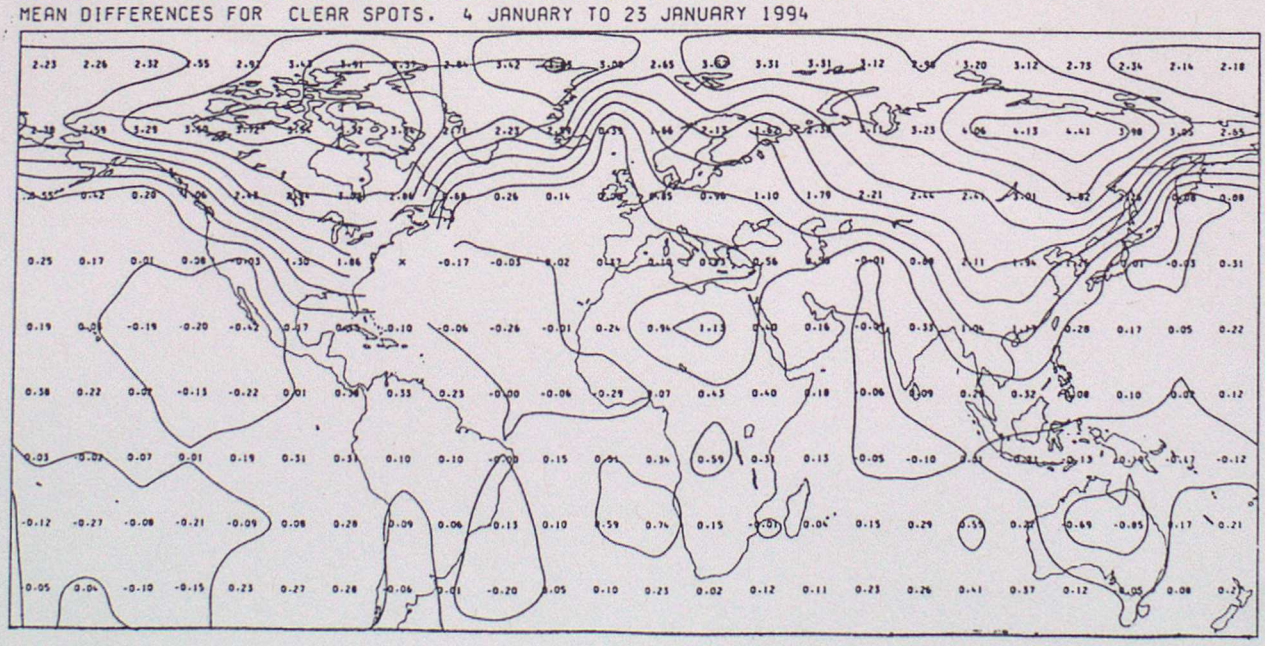
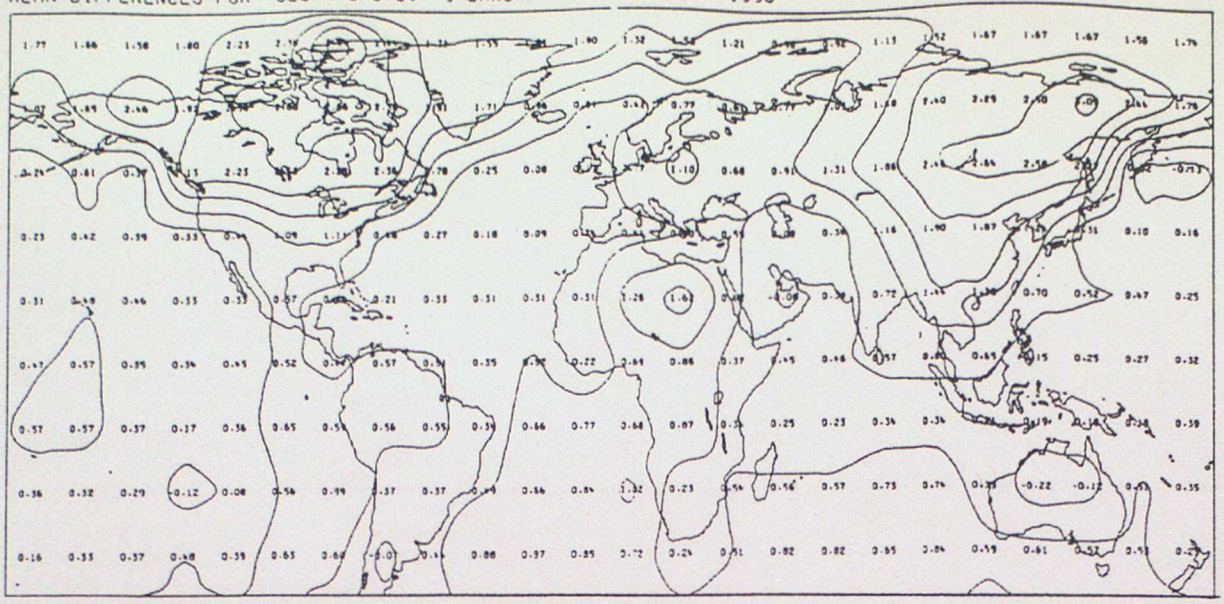
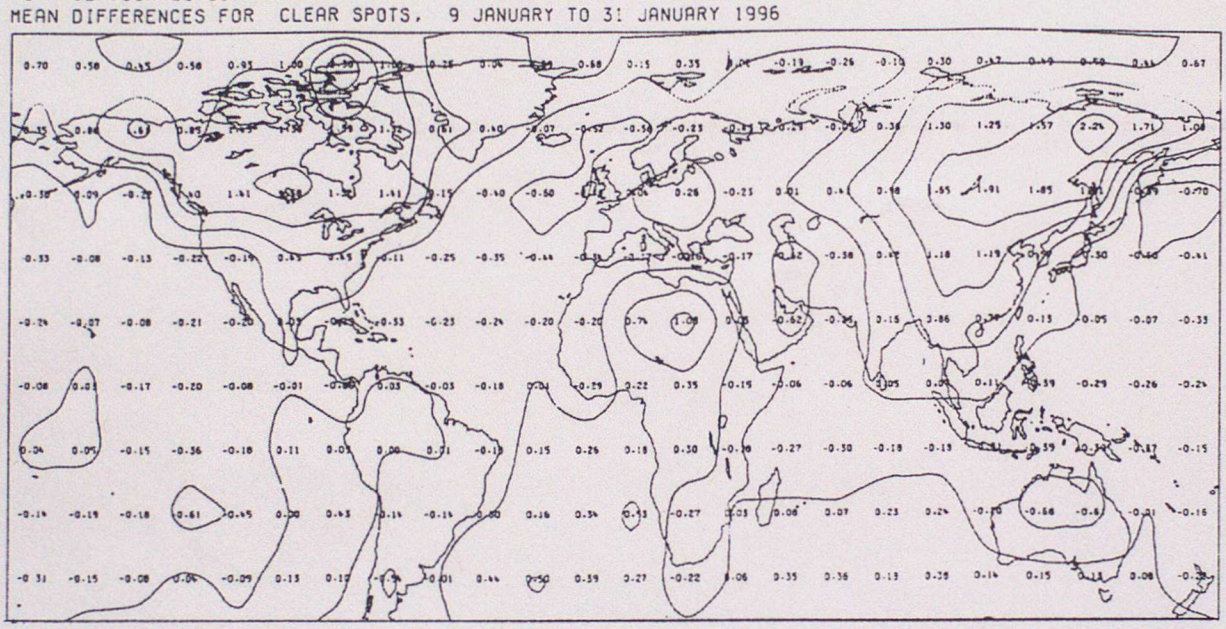


Figure 4.1. Differences between observation and background for NOAA-12 MSU-2 in January 1994 (a) before and (b) after bias correction using the version of GLOSS operational at the time.

(a) NOAA-12 (OBSERVED-BKG) BRIGHTNESS TEMPERATURE DIFFERENCES FOR MSU 2
MEAN DIFFERENCES FOR CLEAR SPOTS, 9 JANUARY TO 31 JANUARY 1996



(b) NOAA-12 (CORRECTED-BKG) BRIGHTNESS TEMPERATURE DIFFERENCES FOR MSU 2
MEAN DIFFERENCES FOR CLEAR SPOTS, 9 JANUARY TO 31 JANUARY 1996



(c) NOAA-12 BIAS CORRECTIONS FOR MSU 2
MEAN DIFFERENCES FOR CLEAR SPOTS, 9 JANUARY TO 31 JANUARY 1996

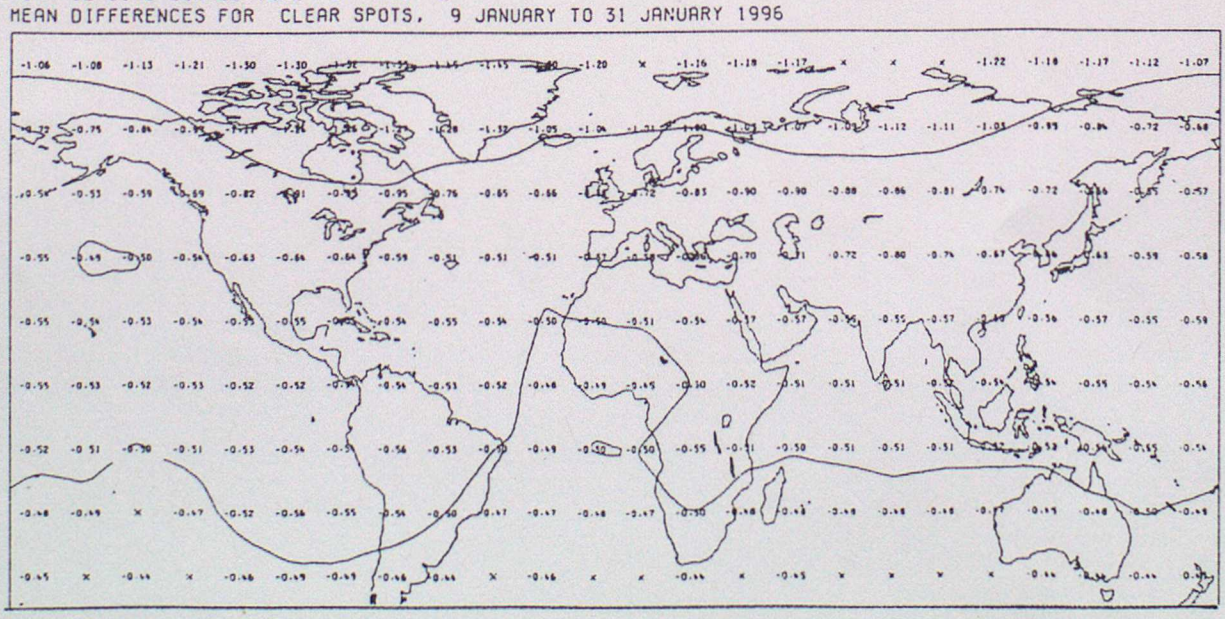


Figure 4.2 Differences between observation and background for NOAA-12 MSU-2 in January 1996 (a) before and (b) after bias correction using the version of GLOSS used in the parallel trial. The magnitude of the bias correction is shown in (c).